



Universidad Politécnica de Madrid
Escuela Técnica Superior de Ingenieros Informáticos



Grupo de Investigación en Ingeniería del Software Empírica

Trabajo Fin de Máster
Máster Universitario en Software y Sistemas

Análisis Post-Hoc de un Experimento acerca de TDD

Autor:
Darwin P. Riofrío Vásquez

Tutores:
Oscar Dieste
Natalia Juristo

Índice general

CAPÍTULO 1. INTRODUCCIÓN	1
1.1. Planteamiento del problema.....	2
1.2. Objetivos	3
1.3. Metodología	3
1.4. Resultados obtenidos.....	4
1.5. Estructura del documento	4
CAPÍTULO 2. ANTECEDENTES.....	6
2.1. Descripción del experimento.....	6
2.1.1. Variables respuesta	7
2.1.2. Factor.....	7
2.1.3. Hipótesis.....	7
2.1.4. Diseño.....	7
2.1.5. Ejecución	8
2.2. Análisis de los datos	8
2.3. Variables demográficas.....	9
CAPÍTULO 3. OBJETIVOS	11
CAPÍTULO 4. APROXIMACIÓN UTILIZADA PARA EL ANÁLISIS POST-HOC.....	13
4.1. Adquisición de Datos Demográficos	13
4.1.1. Cuantificación de los datos.....	14
4.1.2. Dicotomización de datos cuantitativos.....	16
4.2. Análisis estadísticos.....	18
4.2.1. T-Student para muestras independientes	18
4.2.2. U de Mann-Whitney para muestras independientes.....	19
4.2.3. Coeficiente de correlación de Pearson.....	19
4.2.4. Regresión lineal múltiple	19
4.2.5. Stepwise regression	20
CAPÍTULO 5. RESULTADOS	22
5.1. Análisis con datos dicotomizados.....	22
5.1.1. Análisis dicotómico cuando hay normalidad.....	23
5.1.2. Análisis dicotómico cuando NO hay normalidad	25
5.1.3. Resumen de resultados	26
5.2. Análisis de correlación de Pearson.....	28
5.2.1. Variables correlacionadas	28
5.2.2. Variables NO correlacionadas.....	29
5.2.3. Resumen de resultados	31
5.3. Análisis de regresión lineal múltiple.....	32
5.4. Stepwise regression (paso a paso hacia atrás)	34
5.4.1. Influencia de las variables demográficas sobre la Productividad.....	34
5.4.2. Influencia de las variables demográficas sobre la Calidad	38
5.4.3. Influencia de las variables demográficas sobre TUS.....	42
5.4.4. Resumen de resultados	46

CAPÍTULO 6. DISCUSIÓN.....	48
6.1. Dicotomización vs Correlación de Pearson.....	48
6.1.1. Resumen de similitudes y diferencias	49
6.1.2. Dicotomización significativa vs Correlación NO significativa	49
6.1.3. Correlación significativa vs Dicotomización NO significativa	50
6.1.4. Sobre los resultados.....	51
6.2. Regresión lineal vs Dicotomización	52
6.3. Regresión lineal vs Correlación de Pearson	52
6.4. Regresión lineal: entrada forzosa vs stepwise regression.....	53
6.5. Confiabilidad de los resultados obtenidos en los análisis	54
CAPÍTULO 7. CONCLUSIONES	55
REFERENCIAS.....	57
ANEXO A. ANÁLISIS DEMOGRÁFICO.....	58
ANEXO B. ANÁLISIS DICOTÓMICO	97
ANEXO C. ANÁLISIS DE CORRELACIÓN DE PEARSON	243
ANEXO D. CORRELACIÓN VS DICOTOMIZACIÓN	249
ANEXO E. FORMULARIO DEMOGRÁFICO.	278

RESUMEN

Antecedentes: La Ingeniería del Software Experimental (ISE) intenta identificar las variables que influyen en el desarrollo de software. A través de investigación formal identifica las relaciones entre dichas variables y los efectos que producen los distintos valores de las mismas. En los experimentos de Ingeniería del Software (IS), uno de los aspectos más importantes son los sujetos de experimentación. Estos sujetos presentan diferencias notables en cuando a habilidades o experiencia en el desarrollo de software. Esto hace que los resultados de los experimentos se puedan ver influenciados por el tipo de sujetos participantes.

Objetivo: Este Trabajo Fin de Máster (TFM) analiza la importancia de la información poblacional en el análisis de datos de experimentos en IS. Por otro lado se intenta abordar un estudio sobre las posibles metodologías que se pueden usar para realizar dicho análisis.

Contexto: El análisis se realiza sobre los resultados de un experimento llevado a cabo por el Grupo de Investigación de IS empírica de la UPM, cuyo objetivo es analizar el impacto que tiene el uso de TDD (Test Driven Development) sobre la calidad y la productividad del desarrollo software en comparación con el desarrollo tradicional TLD.

Método de Investigación: Se analizan ocho variables demográficas frente a tres variables respuesta. Las metodologías o técnicas de análisis que se revisan son la Dicotomización, la Correlación de Pearson, la regresión lineal múltiple y la stepwise regression.

Resultados: No se encuentran evidencias claras para afirmar que las variables demográficas influyen en los resultados de los experimentos. No obstante los resultados no son del todo concluyentes y queda abierta la investigación a realizarse con una muestra más amplia y representativa. En relación a la metodología de análisis aplicada, la dicotomización y la correlación de Pearson presentan deficiencias que se solventan con la regresión lineal múltiple y la stepwise regression.

Conclusión: Resulta de vital importancia encontrar evidencias de la influencia de las características demográficas de los sujetos experimentales en el análisis de los datos experimentos en IS. Se ha encontrado un buen método para analizar esta influencia, pero falta replicar este análisis a más experimentos de IS para obtener resultados mejor fundados.

ABSTRACT

Background: The Experimental Software Engineering (ESE) attempts to identify the variables that influence the software development. Through formal investigation identifies relationships between these variables and the effects that different values of these variables could produce. In Software Engineering (SE), one of the most important aspects are the experimental subjects. These subjects show remarkable differences in skills or experience in software development. This makes that the results of the experiments can be influenced by the type of participating subjects.

Objective: This Master's Thesis (TFM) discusses the importance of demographic data in the analysis of data from experiments in SE. On the other hand, it attempts to address a study of the possible methodologies that can be used to perform the analysis.

Context: The analysis is performed on the results of an experiment conducted by the ESE Research Group of the UPM, aimed at analyzing the impact of the use of TDD (Test Driven Development) on quality and productivity, compared to traditional development TLD (Test Last Development).

Research Method: Eight demographic variables were analyzed against three response variables. The methodologies and analysis techniques that are reviewed include dichotomization, Pearson correlation, multiple linear regression and stepwise regression.

Results: There is not clear evidence to say that demographic variables influence the results of SE experiments. However the results are not conclusive, and are open to research with a broader and more representative sample. Regarding the applied analysis methodology, dichotomization and Pearson correlation have deficiencies that are solved with multiple linear regression and stepwise regression.

Conclusion: It is very important to find evidence on the influence of demographic characteristics of subjects in the data analysis of SE experiments. We found a good way to analyze this influence, but is necessary replicate this analysis on more SE experiments to obtain sound results.

CAPÍTULO 1. INTRODUCCIÓN

La ingeniería del software es una disciplina que no cuenta con una teoría clara de cómo desarrollar sus actividades. La mayoría de las ingenierías tienen métodos y técnicas para realizar una determinada actividad dentro de sus respectivos campos de aplicación, sin embargo, en la Ingeniería del Software no hay reglas ni modelos claros, así como tampoco procedimientos universalmente aceptados para valorar un modelo frente a otro.

El software que se construye en la actualidad es complejo. Para su desarrollo se utilizan métodos y técnicas cuya efectividad, en general no ha sido demostrada o probada, simplemente se tienen sospechas de que funcionan. Actualmente el software se construye en base a creencias que no tienen una base científica.

Con el fin de proporcionar métodos y técnicas sobre la construcción de software, el campo de la Ingeniería del Software Experimental (ISE) ejecuta experimentos controlados para poder encontrar evidencias de que los modelos empíricos que se utilizan efectivamente funcionan [1].

La experimentación se ha utilizado en otras ciencias con el fin de transformar creencias y sospechas en conocimiento científico. La ISE pretende generar evidencias a través de la experimentación, como se hace en la física, química o la medicina.

En este sentido la ISE persigue identificar las variables que entran en juego cuando se construye software, identificando las relaciones que existen entre dichas variables. Para ello realiza una investigación formal de los efectos que pueden producir los valores que toman esas variables. La investigación experimental hace uso de diseños experimentales para planificar y ejecutar los experimentos y hace uso de la estadística para analizar los resultados.

Por tanto un experimento es una investigación formal que se realiza de forma controlada y rigurosa en la que se manipulan variables de las que se intenta obtener información. Lo que se intenta averiguar es el efecto que tienen los distintos valores que pueden tomar esas variables en el desarrollo de software. Algunas de esas variables son: métodos de diseño, técnicas de pruebas, experiencia de los desarrolladores, etc. Algunos efectos son: calidad, productividad, etc. [1].

En muchas ocasiones es necesario replicar un experimento, es decir, realizar varias ejecuciones de un mismo experimento con el objetivo de tener un tamaño muestral consistente. Los resultados obtenidos en varias repeticiones se combinan con el fin de aumentar la fiabilidad de los resultados y por ende generar evidencias empíricamente fundadas.

Las personas que aplican los métodos o técnicas en la ejecución de los experimentos de IS se les denominan sujetos experimentales. Los sujetos experimentales generalmente presentan diferencias en cuanto a habilidades o experiencia. Este TFM analiza la influencia que puede tener sobre los resultados de un experimento, el tener en cuenta la información poblacional, es decir, las condiciones particulares de los sujetos que participan en el experimento

La línea general en la que se encuadra este Trabajo Fin de Máster (TFM) es la generación de evidencias a partir de la combinación de resultados de varios experimentos. Más específicamente se encuadra dentro del proyecto FIDIPRO del que forma parte el grupo de investigación en Ingeniería del Software Experimental de la Universidad Politécnica de Madrid (UPM). Se trata de un proyecto de investigación que tiene como objetivo crear un laboratorio industrial de ingeniería del software experimental (ESEIL). ESEIL ejecuta experimentos controlados para reunir conocimientos sobre el comportamiento de los diferentes métodos y técnicas en diferentes entornos. Los resultados experimentales que se obtienen conducirán a conclusiones acerca de las tecnologías de software anteriores y sus contextos de aplicación. Además ayudará a los organizadores a aprender más acerca de las variables que influyen en su proceso de desarrollo de software y el control de las variables identificadas para mejorar sus procesos y hacerlos más previsibles.

1.1. Planteamiento del problema

Los sujetos experimentales que participan en los experimentos de ingeniería del software no son ni mucho menos similares. Al contrario, presentan diferencias notables en cuanto a habilidades o experiencia en el desarrollo de software [2]. Esto hace que los resultados de los experimentos se puedan ver influenciados en muchos de los casos por el tipo de sujetos participantes. En consecuencia, surge la necesidad de identificar las

variables relacionadas con las cualidades y habilidades de los sujetos experimentales (variables demográficas), que presumiblemente pueden influir en los resultados de los experimentos que lleva a cabo la ISE.

Resulta de suma importancia realizar este estudio dado el impacto negativo que puede tener la influencia de las variables demográficas sobre los resultados de los experimentos de la ISE. Tener en cuenta la información demográfica ayudará a obtener resultados más fiables y que derivarán en conclusiones correctas acerca de los factores que se estén analizando. Las conclusiones a las que se llega ayudará a los profesionales del software a construir software de mejor calidad y más competitivo.

1.2. Objetivos

El objetivo de este TFM es doble. Por un lado se pretende analizar la posible influencia que puede tener hacer uso de la información demográfica de los sujetos experimentales, al realizar el análisis de los datos que se obtienen como resultado de la ejecución de los experimentos por parte de la ISE. Por otro lado, se intenta determinar las bondades y puntos flacos de las distintas metodologías que se pueden usar para realizar dicho análisis. Para alcanzar estos objetivos se realiza un análisis Post-Hoc de un experimento llevado a cabo por el grupo de investigación en Ingeniería del Software Experimental de la UPM en el ámbito de FIDIPRO.

1.3. Metodología

Lo primero que se hace es construir los conjuntos de datos sobre los que se va a realizar el análisis. Las variables demográficas se obtienen a través de formularios demográficos que rellenan los sujetos antes o después de la ejecución de un experimento.

Sobre los datos obtenidos a través de los formularios demográficos se construyen dos conjuntos de datos, uno con datos tipo escala y otro con datos dicotomizados. El proceso de dicotomización consiste en separar en dos grupos los valores de cada variable demográfica. Sobre estos conjuntos de datos se realiza diferentes tipos de análisis.

Para analizar su influencia se construye un conjunto de datos que se fusiona con los datos que se recogen durante el experimento. De esta fusión resulta un conjunto de datos más completo y que incluye los factores, las variables respuesta y las variables demográficas.

Sobre este conjunto de datos se realiza varios análisis estadísticos con el fin de encontrar evidencias de que existe influencia de las variables demográficas sobre los resultados obtenidos en la ejecución del experimento. Las metodologías o técnicas de análisis que se revisan son la dicotomización, la correlación de Pearson, la regresión lineal múltiple y la stepwise regression.

1.4. Resultados obtenidos

Una vez realizados las diferentes técnicas de análisis de los datos y tras estudiar los resultados obtenidos, en términos generales se puede concluir que no se han encontrado evidencias suficientes como para afirmar que las variables demográficas influyan sustancialmente en los resultados del experimento. Sin embargo se encuentran resultados puntuales en los que parece que hay indicios de dicha influencia, sin embargo no se han considerado suficientes como para realizar dicha afirmación.

Con respecto a las diferentes técnicas de análisis de los datos, se podrá ver que realizar una dicotomización ayuda a comprender los resultados de una forma más intuitiva para el ser humano, pero que no es una buena opción. En el proceso de dicotomización se puede perder mucha información relevante. Mejorando las deficiencias que tiene la dicotomización se presenta la correlación de Pearson como una opción de mejora, aunque aún no resulta ser la mejor opción, dado que el análisis de las variables se realiza por pares, sin tener en cuenta las demás variables. Veremos que esta deficiencia la solventa la stepwise regression.

1.5. Estructura del documento

En el capítulo 2: Antecedentes, se presentan los acontecimientos previos al trabajo que se ha realizado, es decir, se cuenta el trabajo más general en el que se enmarca este TFM. Dado que este trabajo precede a un experimento de ingeniería del software, se realiza una descripción de dicho experimento.

En el capítulo 3: Objetivos, se presentan los objetivos del TFM, se enumeran los objetivos específicos como los sub-objetivos que derivan de los mismos.

El capítulo 4: Aproximación utilizada para el análisis Post-Hoc, contiene el procedimiento que se ha llevado a cabo para el análisis de los datos demográficos, se comenta el proceso de dicotomización, así como los métodos estadísticos que se han usado en el análisis de los datos.

El capítulo 5: Resultados, recoge un resumen de todo el trabajo realizado y que se encuentra detallado en los anexos. Se presentan los cuatro tipos de análisis que se han llevado a cabo: Dicotomización, Correlación de Pearson, Análisis de regresión lineal múltiple y stepwise regression.

En el capítulo 6: Discusión, se hace una reflexión sobre los resultados obtenidos y se realiza un estudio comparativo de los tres tipos de análisis que se han llevado a cabo, comentando las diferencias encontradas, así como las deficiencias que se han ido encontrando a los diferentes tipos de análisis.

El capítulo 7: Conclusión, contiene una síntesis del trabajo realizado, y se intenta dar respuesta a los objetivos planteados.

CAPÍTULO 2. ANTECEDENTES

No está claro si el desarrollo usando TDD (Test Driven Development) mejora la productividad y la calidad del software frente al desarrollo usando TLD (Test Last Development). Por este motivo el Grupo de Investigación de la UPM ha realizado experimentos sobre el impacto del desarrollo usando TDD en comparación con el desarrollo tradicional TLD. El objetivo era evaluar la *Calidad* y la *Productividad* que se obtiene usando cada técnica, determinando bajo qué circunstancias es mejor utilizar una u otra. Estos experimentos se han realizado en el mundo académico.

La mayoría de los experimentos acerca de TDD se han desarrollado en el ámbito académico, hay muy pocos experimentos en la industria, con lo cual sus resultados no pueden extrapolarse directamente al mundo profesional.

El proyecto ESEIL proporciona la oportunidad de experimentar en la industria. Por lo que el grupo de investigación en ISE de la UPM ha realizado experimentos en el entorno profesional. Por lo tanto el contexto en el que se desarrolla este TFM, es un experimento llevado a cabo por el grupo de investigación en ISE de la UPM, cuyo objetivo es analizar el impacto que tiene el uso de TDD sobre la Calidad y la Productividad del desarrollo software en comparación con el desarrollo tradicional TLD, desde el punto de vista de los profesionales del desarrollo de software. A continuación se describe el experimento sobre el que se va a realizar el análisis Post-Hoc.

2.1. Descripción del experimento

Este apartado contiene la descripción del experimento que se ha realizado previo a este análisis Post-Hoc. Se describen cuales son la variables del experimento, los factores que se evalúan, las hipótesis planteadas y el diseño experimental que se uso para su ejecución.

2.1.1. Variables respuesta

La efectividad de un enfoque de programación se puede evaluar desde diferentes perspectivas. Experimentos anteriores a este han utilizado variedad de métricas. Sin embargo en este experimento se usan la calidad y la productividad, puesto que son las que se han usado más habitualmente en experimentos previos, lo que permitirá adicionalmente comparar los resultados obtenidos.

Por lo tanto las variables respuesta del experimento son: *Productividad, Calidad y TUS (volumen de trabajo)*.

2.1.2. Factor

Las variables independientes del experimento son las técnicas de desarrollo: Test Driven Development (TDD) y Test Last Development (TLD). Tanto en TDD como en TLD, se realizan las siguientes tareas: descomposición en sub-tareas, codificación, pruebas y refactorización. La única diferencia es que en TLD las pruebas se realizan después de escribir el código, mientras que TDD es un enfoque de desarrollo evolutivo que combina Test First Development (TFD) y las refactorización. El desarrollo TFD consiste en escribir primero una prueba y a continuación el código que la satisface.

Dado que se realizan dos replicas del en las que se realizan las tareas usando la técnica TDD, el experimento cuenta con un factor (técnica de desarrollo) con tres niveles. Los tratamientos son TLD, TDD1 y TDD2.

2.1.3. Hipótesis

Las hipótesis del experimento son:

- H_{10} : La Calidad es la misma en TDD y TLD
- H_{11} : La Calidad es diferente TDD y TLD
- H_{20} : La Productividad es la misma en TDD y TLD
- H_{21} : La Productividad es diferente en TDD y TLD

2.1.4. Diseño

El diseño del experimento es el que se muestra en la Tabla 2.1.

Grupo	Secuencia temporal				
	Entrenamiento	Sesión	Entrenamiento	Sesión	Sesión
G1	TLD	TLD	TDD	TDD1	TDD2

Tabla 2.1. Diseño experimental

Primero se da un entrenamiento del tratamiento TLD, luego se realiza una sesión experimental en la que los sujetos programan usando la estrategia TLD. A continuación a

los sujetos se les da entrenamiento en TDD y finalmente se realizan dos sesiones experimentales TDD1 y TDD2, en las que los sujetos programan usando la técnica TDD.

Para la relación de cada sesión experimental se proporciona a los sujetos una tarea experimental diferente: Mars Rover en caso de TLD, Bowling Scorekeeper en el caso de TDD1 y MusicPhone en el caso de TDD2. Bowling Scorekeeper es una tarea “greenfield” que consiste en desarrollar a partir de cero. Mientras que MusicPhone es una tarea “legacy code” que consiste en desarrollar a partir de código existente, por ejemplo con el fin de dotarle de nuevas funcionalidades o porque se quiere añadir compatibilidad con otro tipo de hardware.

2.1.5. Ejecución

El experimento se ha replicado en tres sedes de en una empresa finlandesa, por confidencialidad no se proporcionará el nombre de la misma. Además también por confidencialidad se mantienen ocultos los datos crudos, esto es, los datos obtenidos en el experimento.

Al empezar se repartió un protocolo para que los sujetos tuvieran conocimientos de las condiciones de realización de cada tarea, así como del tiempo del que disponían para la realización de la misma.

Una vez terminado el tiempo para la realización de la tarea correspondiente, los investigadores realizaron la recolección de los datos en el lugar donde se realizó el experimento. Se realizaron copias de las carpetas que contenían el código fuente en una unidad de almacenamiento. Para la recolección de los datos demográficos se utilizaron formularios en línea. Estos formularios en línea fueron volcados automáticamente a hojas de cálculo.

2.2. Análisis de los datos

Tras la realización del experimento se lleva a cabo el análisis de los datos obtenidos, para realizar el contraste de las hipótesis planteadas en el diseño del experimento. Se realizan varios análisis estadísticos para testear las hipótesis planteadas: T de Student para muestras relacionadas, el test no paramétrico Mann-Whitney y el Modelo Lineal General para medidas repetidas (MLG). Sin embargo no se profundizara más en el detalle de este análisis dado que no es relevante para este TFM.

Los resultados del análisis de los datos indican que no hay diferencias entre TLD y TDD. TDD1 resulta superior a TLD y TDD2. Hay indicios de que los resultados que se obtienen se deben en parte al carácter “greenfield” de la tarea Bowling Scorekeeper.

2.3. Variables demográficas

Los sujetos experimentales son uno de los componentes más importantes durante la realización de experimentos de ingeniería del software. Como se ha comentado, los sujetos experimentales generalmente presentan diferencias significativas en cuanto a habilidades y experiencia en el ámbito de la ingeniería del software. Dado que se tiene la sospecha de que las cualidades particulares de un sujeto pueden influir en los resultados obtenidos en experimentos de ingeniería del software, tal y como se explica en el apartado 5.1.5, tras la realización del experimento los sujetos experimentales rellenaron los formularios demográficos, con el fin de obtener información en detalle sobre sus cualidades y habilidades.

Las preguntas que se realizó a los sujetos tienen que ver con su experiencia académica, experiencia profesional, experiencia en programación, experiencia en pruebas de unidad, experiencia en metodologías de desarrollo, experiencia en el uso de IDEs y experiencia en TDD. Con los datos obtenidos en los formularios demográficos se realizará un análisis Post-Hoc que es objeto de esta tesis. En el Anexo E. Formulario demográfico, se encuentra el detalle de las preguntas a las que respondieron los sujetos experimentales. La Figura 2.1 muestra un fragmento de dicho formulario. La información recogida en el formulario demográfico puede resumirse en un conjunto de variables demográficas con las que se realizará el análisis Post-Hoc. Estas variables son las siguientes:

- Experiencia profesional y académica
- Experiencia de programación
- Experiencia con el lenguaje de programación Java
- Experiencia en pruebas de unidad
- Experiencia en metodologías de desarrollo de software
- Experiencia en entornos de desarrollo integrado (IDEs)
- Experiencia en TDD

8. What IDE (Integrated Development Environment) do you currently use? *

9. How would you rate your experience in Java? *

- ☐ No experience (<2 years)
- ☐ Novice (2-<=5 years)
- ☐ Intermediate (5-<=10 years)
- ☐ Expert (>10 years)

10. How would you rate your experience with Junit testing framework? *

- ☐ No experience (<2 years)
- ☐ Novice (2-<=5 years)
- ☐ Intermediate (5-<=10 years)
- ☐ Expert (>10 years)

11. Have you ever used TDD as a development methodology? *

- ☐ Yes
- ☐ No

12. If "Yes" above, how would you rate your experience with TDD?

- ☐ No experience (<2 years)
- ☐ Novice (2-<=5 years)
- ☐ Intermediate (5-<=10 years)
- ☐ Expert (>10 years)

Figura 1. Fragmento del formulario demográfico

CAPÍTULO 3. OBJETIVOS

El propósito del presente trabajo es realizar un análisis Post-Hoc del experimento descrito en el capítulo 2. Este análisis consiste en averiguar la influencia que pueden ejercer las variables demográficas, tales como la experiencia de los sujetos experimentales, en la efectividad de las estrategias de programación TDD y TLD.

Se pueden plantear dos objetivos principales. El primero consiste a analizar la influencia de las variables demográficas sobre las variables respuesta del experimento.

Como se mencionó en el capítulo anterior, se analizan varias variables que tienen que ver con la experiencia o habilidad de los sujetos experimentales, por lo tanto el primer objetivo puede descomponerse en los siguientes sub-objetivos:

- Analizar la influencia de la *Experiencia Profesional y Académica* sobre la *Calidad* y la *Productividad* observada en el experimento para los estilos de programación TLD y TDD.
- Analizar la influencia de la *Experiencia de Programación* sobre la *Calidad* y *Productividad* observada en el experimento para los estilos de programación TLD y TDD.
- Analizar la influencia de la *Experiencia en Java* sobre la *Calidad* y la *Productividad* observada en el experimento para los estilos de programación TLD y TDD.
- Analizar la influencia de la *Experiencia en Pruebas de Unidad* sobre la *Calidad* y la *Productividad* observada en el experimento para los estilos de programación TLD y TDD.

- Analizar la influencia de la *Experiencia en Metodologías de Desarrollo* sobre la *Calidad* y la *Productividad* observada en el experimento para los estilos de programación TLD y TDD.
- Analizar la influencia de la *Experiencia en Entornos de Desarrollo integrado* (IDEs) sobre la *Calidad* y la *Productividad* observada en el experimento para los estilos de programación TLD y TDD.
- Analizar la influencia de la *Experiencia en TDD* sobre la *Calidad* y la *Productividad* observada en el experimento para los estilos de programación TLD y TDD.

Por otro lado, un análisis Post-Hoc puede realizarse utilizando diversas técnicas estadísticas y de análisis de los datos. Sin embargo no todas ellas proporcionan los mismos resultados. Por esta razón surge la necesidad de plantear un segundo objetivo, que consiste en analizar la influencia que tiene el método de análisis en los resultados que se obtienen al realizar el primer objetivo.

CAPÍTULO 4. APROXIMACIÓN UTILIZADA PARA EL ANÁLISIS POST-HOC

El presente capítulo contiene una descripción del proceso de análisis que se ha llevado a cabo. Se presentan los métodos aplicados tanto para la obtención de los datos como para la realización de los análisis estadísticos relevantes.

4.1. Adquisición de Datos Demográficos

La primera tarea que se ha realizado es la adquisición de los datos demográficos. Consiste en examinar las respuestas de los formularios que recogen la información personal de los sujetos para obtener los datos que se utilizan para el análisis. La información recogida en los formularios demográficos se utiliza para realizar un proceso de selección de posibles variables moderadoras del experimento. Para que una variable sea considerada posible moderadora del experimento se tiene en cuenta que haya diferencias sustanciales entre lo que responden los sujetos. Por ejemplo se ha preguntado por el entrenamiento en TDD, pero ningún sujeto aporta información sobre ello. Entonces esta variable no es tomada en cuenta como posible variable moderadora. Otro caso en el que no es tomada en cuenta es cuando la respuesta es la misma para todos los sujetos. Tras un exhaustivo análisis se identificaron las siguientes variables:

- Experiencia Académica y Profesional
- Experiencia en Programación
- Experiencia en Java
- Experiencia en Testing
- Experiencia en JUnit
- Experiencia en el uso de IDE

- Experiencia en Metodologías de Desarrollo
- Experiencia en TDD

Las variables seleccionadas se han codificado de dos formas distintas, uno tiene datos cuantitativos (esto es, las variables poseen escalas de tipo intervalo o incluso de tipo ratio). Y mediante datos dicotómicos (esto es, las variables solo poseen dos valores, que denominaremos “-” y “+”). La cuantificación de los datos se explica en la sección 4.1.1. El proceso de dicotomización se explica en la sección 4.1.2. El detalle completo del análisis demográfico se encuentra en el Anexo A. Análisis demográfico.

4.1.1. Cuantificación de los datos

Las variables recogidas a través de los formularios demográficos son de diferentes tipos. Algunos son binarios, como cuando se pregunta por si han usado Katas o no. Otros son de tipo escala, como cuando se pregunta por los años de experiencia en programación. Finalmente, en ocasiones las variables son de tipo intervalo, como cuando se pregunta dando opciones como: sin experiencia, novato, intermedio, experto. Dependiendo de cada situación se toma un procedimiento diferente para cuantificar los datos.

- En el caso de los datos binarios como es el caso de la variable *experiencia en IDE*, se ha tomado un valor “0” cuando la respuesta es No, y un “1” cuando la respuesta es Sí.
- Para el caso de los datos numéricos como es el caso de la variable *experiencia profesional y académica*, no hay problema porque la traducción es directa; se corresponde con los años de experiencia que poseen los sujetos.
- Cuando los datos se encuentran en escala de intervalo, a la hora de cuantificarlos nos quedamos con el valor del límite inferior del intervalo, Por ejemplo cuando se pregunta por la *experiencia en Java* se presentan las siguientes opciones:
 - Sin experiencia (0 – 2 años)
 - Novato (2 - 5 años)
 - Intermedio (5 – 10 años)
 - Experto (más de 10)

Por tanto en este caso los datos se cuantifican de esta forma: Sin experiencia se toma el valor “0”, Novato se toma el valor “2”, Intermedio se toma el valor “5” y para los expertos se toma el valor “10”.

Realizando este procedimiento para cada una de las variables demográficas y para cada uno de los experimentos se obtienen los datos de la tabla 4.1. La primera columna de la tabla muestra el identificador (ID) del sujeto experimental. Las demás columnas se corresponden con los valores cuantitativos de las variables demográficas.

ID	Académica y profesional	Programación	Java	Testing	Junit	IDE	Metodologías de desarrollo	TDD
S01	14	10	5	2	2	1	1	0
S02	6	10	2	5	2	0	1	1
S03	30	10	10	10	2	1	1	0
S04	10	10	5	5	5	1	1	2
S05	7,5	5	5	5	5	1	1	2
S06	0	2	0	2	0	1	1	0
S07	10	5	2	2	0	0	1	0
FSH01	0	2	0	0	0	0	0	1
FSH02	15	10	10	10	10	1	1	0
FSH03	14	5	2	5	0	0	1	2
FSH04	9	5	0	2	0	1	1	1
FSH05	3	5	2	2	2	0	0	1
FSH11	0	10	0	0	0	1	1	0
FSH13	15	10	5	2	2	1	1	0
FSH14	6	10	5	2	5	0	1	1
FSK01	6	5	0	2	0	1	1	1
FSK02	6	5	0	2	0	1	1	0
FSK03	0,25	2	2	0	0	1	0	0
FSK04	9	5	2	2	2	1	1	2
FSK05	5	5	2	2	2	1	1	0
FSK06	0	2	0	0	0	1	0	0
FSK07	5	5	2	2	0	1	0	1
FSK08	10	5	2	0	0	1	1	1
FSK09	5	5	2	2	2	1	1	2
FSK11	5	5	0	0	0	1	0	0
FSK12	11	10	10	2	2	1	1	2
FSK13	5	5	2	2	2	0	1	0
FSK14	9	5	0	2	0	0	1	0
FSK16	15	10	2	0	0	1	1	0
FSK18	9	2	0	2	0	1	1	1
FSK19	0	5	2	0	0	0	1	0

Tabla 4.1. Datos cuantitativos

4.1.2. Dicotomización de datos cuantitativos

El proceso de dicotomización consiste en clasificar en dos grupos una serie de datos cuantitativos. Por ejemplo, si tenemos una serie de datos que representan las notas de un grupo de alumnos, podemos realizar una simplificación quedándonos con dos categorías que pueden ser aprobados y suspensos. Generalmente este tipo de estrategia se usa como preparación de datos para un posterior análisis. Se presume que la motivación para utilizar este tipo de técnica es facilitar la comprensión de los datos, puesto que para el ser humano es más cómodo asimilar información del tipo ganancias-perdidas, nuevo-viejo etc.

La dicotomización se usa en el presente trabajo con el objetivo de clasificar la experiencia de los sujetos que participan en el experimento en dos categorías, los que se presume que pueden influir negativamente, y los que pueden influir positivamente en los resultados de los experimentos llevados a cabo. Por tanto se realizó un análisis los datos cuantitativos para generar a partir de ellos los datos dicotómicos.

Un ejemplo del proceso de dicotomización de los datos cuantitativos los podemos ver con la experiencia de programación en Java. Los resultados obtenidos en el formulario demográfico son los siguientes:

ID	Java experience
S01	Intermediate (5-<=10 years)
S02	Novice (2-<=5 years)
S03	Expert (>10 years)
S04	Intermediate (5-<=10 years)
S05	Intermediate (5-<=10 years)
S06	No experience (<2 years)
S07	Novice (2-<=5 years)

Tabla 4.2. Respuestas de los sujetos sobre la experiencia en Java

Se considera que el sujeto S06 está en desventaja con respecto al resto de sujetos puesto que es el único que no tiene experiencia, y por tanto presumiblemente puede influir negativamente en los resultados del experimento. De esta forma se realiza una simplificación dividiendo los datos en dos categorías, los que influyen Negativamente (S06) y los que presumiblemente pueden influir de forma Positiva (S01, S02, S03, S04, S05 y S07). Los que pertenecen a la categoría Negativa se marcan con “-” y los que pertenecen a la categoría Positiva se marcan con “+”. En la Tabla 4.3 se muestran los datos obtenidos para todas las variables tras el proceso de dicotomización. La primera columna de la tabla muestra el identificador (ID) del sujeto experimental. Las demás columnas se corresponden con los valores dicotomizados de las variables demográficas.

ID	Académica y profesional	Programación	Java	Testing	Junit	IDE	Metodologías de desarrollo	TDD
S01	+	+	+	-	+	+	+	-
S02	+	+	+	+	+	-	+	+
S03	+	+	+	+	+	+	+	-
S04	+	+	+	+	+	+	+	+
S05	+	+	+	+	+	+	+	+
S06	-	-	-	-	-	+	+	-
S07	+	+	+	-	-	-	+	-
FSH01	-	-	-	-	-	-	-	+
FSH02	+	+	+	+	+	+	+	-
FSH03	+	+	+	+	-	-	+	+
FSH04	+	+	-	+	-	+	+	+
FSH05	+	+	+	+	+	-	-	+
FSH11	-	+	-	-	-	+	+	-
FSH13	+	+	+	+	+	+	+	-
FSH14	+	+	+	+	+	-	+	+
FSK01	+	+	-	+	-	+	+	+
FSK02	+	+	-	+	-	+	+	-
FSK03	-	-	+	-	+	+	-	-
FSK04	+	+	+	+	+	+	+	+
FSK05	+	+	+	+	+	+	+	-
FSK06	-	-	-	-	-	+	-	-
FSK07	+	+	+	+	-	+	-	+
FSK08	+	+	+	-	+	+	+	+
FSK09	+	+	+	+	+	+	+	+
FSK11	+	+	-	-	-	+	-	-
FSK12	+	+	+	+	+	+	+	+
FSK13	+	+	+	+	+	-	+	-
FSK14	+	+	-	+	-	-	+	-
FSK16	+	+	+	-	-	+	+	-
FSK18	+	-	-	+	+	+	+	+
FSK19	-	+	+	-	-	-	+	-

Tabla 4.3. Datos dicotomizados

El proceso completo y detallado para la obtención de los datos dicotómicos se encuentra en el Anexo A. Análisis demográfico.

4.2. Análisis estadísticos

El primer análisis estadístico que se realiza es el de comparación de medias. Se trata de un test estadístico que se aplica al analizar los datos dicotómicos, con el objetivo de comprobar si existe una diferencia significativa de las medias de las dos categorías (Positiva, Negativa). Para realizar esta comprobación es necesario aplicar un test estadístico. Hay varios tipos de test que se pueden aplicar para comparar las medias. Para poder elegir el tipo de test adecuado, lo primero que se hace es comprobar que la distribución de los datos es Normal. En caso afirmativo se aplica el test de comparación de medias de la T de Student y en caso de que los datos no sigan una distribución Normal se aplica un test no paramétrico, en concreto el test de la U de Mann-Whitney.

El segundo análisis que se lleva a cabo es el de correlación de Pearson. En este caso se realiza una correlación bivariada.

El tercer análisis que se realiza es el de regresión lineal múltiple, el análisis se realiza con entrada forzosa, es decir, el modelo se realiza con todas las variables demográficas.

El cuarto análisis que se realiza es la regresión paso a paso hacia atrás (stepwise regression). En este caso se introducen todas las variables independientes con las que se pretende construir el modelo de regresión lineal y se van realizando pasos sucesivos en los que se va quitando la variable que presenta menor significación.

4.2.1. T-Student para muestras independientes

Se trata de un test paramétrica de contraste de hipótesis que se usa para comparar la media de una variable dependiente entre dos grupos independientes de sujetos.

El test de la T de Student para muestras independientes permite realizar un análisis de la relación entre una variable cuantitativa y otra cualitativa. Es uno de los análisis estadísticos más comunes y ampliamente utilizados.

La aplicación del test de la T de Student para muestras independientes es un test que requiere que las categorías de las observaciones cumplan dos condiciones. Por un lado tienen que cumplir que haya normalidad, y por otro lado que haya igualdad de varianzas.

La hipótesis nula (H_0) de este test sostiene que no hay diferencia significativa en las medias de los grupos. Mientras que la hipótesis alternativa (H_1) sostiene que si hay diferencia significativa entre las medias de los grupos. Para comprobar la hipótesis de normalidad se ha usado el test de Shapiro-Wilk, mientras que para comprobar la igualdad de varianzas se utiliza el test de Levene.

4.2.2. U de Mann-Whitney para muestras independientes

Esta prueba es similar a la T de Student para muestras independientes. El test de la U de Mann-Whitney para muestras independientes que permite realizar un análisis de la relación que existe entre una variable cuantitativa y otra cualitativa.

Se trata de un test no paramétrico que se aplica cuando las variables no presentan Normalidad. La hipótesis nula H_0 afirma que las categorías se comportan como dos poblaciones idénticas, es decir, no hay diferencias entre la mediana de cada uno de los grupos. La hipótesis alternativa H_1 afirma que hay diferencias entre las categorías.

En nuestro caso de estudio el test de la U de Mann-Whitney la utilizamos como alternativa a la T de Student cuando la distribución de los datos no es Normal.

4.2.3. Coeficiente de correlación de Pearson

Es un test estadístico de contraste paramétrico de hipótesis, se usa para analizar la relación entre dos variables medidas en un nivel de intervalo o escala. El coeficiente de correlación es un número que toma valor en el intervalo -1 y 1, el signo denota la pendiente de la recta de ajuste lineal, mientras que la magnitud indica el grado de dependencia lineal entre las dos variables. La correlación de Pearson no evalúa causalidad, por lo tanto aunque permite averiguar si una variable está relacionada con otra de algún modo, no indica que una sea consecuencia de la otra.

En la correlación de Pearson la hipótesis nula H_0 indica que las dos variables son independientes, mientras que la hipótesis alternativa H_1 indica que las dos variables en estudio están linealmente relacionadas.

El coeficiente de correlación de Pearson mide el grado de covariación entre dos variables que se suponen relacionadas linealmente. En el caso de variables que se encuentren relacionadas pero no de forma lineal, el índice de correlación de Pearson no se puede aplicar.

La aplicación del test del coeficiente de correlación de Pearson no asume normalidad, por lo tanto no es necesario realizar el test de normalidad a las variables.

Una de las interpretaciones que se le puede dar al resultado de la correlación, que como se mencionó anteriormente es un valor comprendido entre -1 y 1, es la propuesta por Cohen, según el valor absoluto del coeficiente de correlación obtenido, se puede decir que la fuerza de la relación entre las variables es: Baja (0,1), Media (0,3), Alta (0,5).

4.2.4. Regresión lineal múltiple

El modelo de regresión lineal simple expresa la relación entre una variable llamada regresando y otra variable llamada regresor. La variable regresando es la que se pretende

explicar (esto es, la variable dependiente) mientras que el regresor es la variable explicativa del modelo (esto es, la variable independiente). La relación que liga las variables del modelo es de tipo lineal. El modelo de regresión lineal simple se expresa matemáticamente como:

$$Y = a + bX + \varepsilon$$

La regresión lineal permite realizar un análisis entre dos variables cuantitativas, una variable independiente X y una variable dependiente Y . a es el término independiente e indica el valor de la Y cuando el valor de la X es 0. b es el coeficiente de regresión, indica lo que cabe esperar que cambie la respuesta por cada incremento unitario en la X , es decir la pendiente de la recta. El símbolo ε representa el error.

La regresión lineal múltiple es una extensión del modelo de regresión lineal simple, con la diferencia de que en la regresión lineal múltiple hay múltiples variables explicativas. El modelo matemático de la regresión lineal múltiple se representa con la siguiente fórmula:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$

Y es la variable dependiente o regresando, X_i son las variables independientes o regresores, β_i miden la influencia que tiene la variable X_i sobre el modelo de regresión. De la misma forma que en la regresión lineal simple, los coeficientes β_i indican la influencia que tiene el incremento unitario de las X_i en las variables explicativas o independientes. β_0 representa el valor de la Y cuando todas las X_i valen 0.

En la regresión lineal múltiple la hipótesis nula H_0 sostiene que no existe dependencia lineal entre la variable dependiente y las variables independientes, es decir, que los coeficientes de regresión (β_x) son 0. La hipótesis alternativa H_1 sostiene que si existe dicha dependencia. Si se rechaza la hipótesis nula significa que algún coeficiente es significativo. Sin embargo, si se acepta la hipótesis nula, se concluye que no hay evidencia de exista relación lineal entre la variable dependiente y ninguna de las variables independientes.

La aplicación de la regresión lineal exige que los datos cumplan con ciertas condiciones como la homogeneidad de varianzas y la normalidad de residuos. No obstante, por tratarse de aspectos avanzados, apenas tratados en el MUSS, en este TFM no serán tenidos en consideración.

4.2.5. Stepwise regression

La stepwise regression es una variante de la regresión lineal múltiple. Esta variación tiene que ver con el método de selección de variables a introducir en el modelo.

El método estándar de selección de variables independientes es el que se usa en la regresión lineal múltiple, consiste en introducir a la vez, todas las variables independientes

con las que se quiere construir el modelo de predicción. Por el contrario en la stepwise regression, en concreto en la regresión paso a paso hacia atrás, se introducen todas las variables independientes con las que se pretende construir el modelo de regresión lineal y se va realizando pasos sucesivos en los que se va quitando la variable que presenta menor significación.

La stepwise regression proporciona un modelo en cada paso. De los modelos que se obtienen hay que elegir el más adecuado. Una buena estrategia para la elección del modelo de regresión es el número de variables y el coeficiente de determinación R^2 . En ocasiones se puede querer tener un modelo de predicción no muy complicado, por tanto se elige el modelo que tenga menos variables independientes. Por otro lado el coeficiente de determinación R^2 es un factor importante que hay que tener en cuenta, ya que sirve para evaluar la bondad de ajuste o poder explicativo del modelo obtenido en cada paso de la stepwise regression.

El coeficiente de determinación R^2 es un valor que está acotado entre 0 y 1, cuando más se aproxime a 1, mayor bondad de ajuste o poder explicativo presenta modelo.

CAPÍTULO 5. RESULTADOS

Este apartado presenta un resumen de los resultados de los análisis que se han obtenido siguiendo los métodos que se han descrito en el capítulo 4.

5.1. Análisis con datos dicotomizados

El análisis dicotómico que se ha llevado a cabo para averiguar si existe relación entre las variables demográficas (ej. experiencia de programación, uso de TDD, etc.) y las variables respuesta del experimento (ej. *Productividad*, *Calidad*, etc.).

Para realizar este análisis se han usado los datos dicotomizados que se obtienen tras la recolección de datos demográficos, descritos en el capítulo anterior. El análisis dicotómico examina cada una de las variables demográficas frente a cada una de las variables dependientes obtenidas en la realización de cada uno de los experimentos. De esta forma, al tener tres variables dependientes (TUS, Quality, Productivity), ocho variables demográficas y tres experimentos, se han analizado $3 \times 3 \times 8 = 72$ pares de variables.

El análisis con datos dicotómicos consiste en aplicar un test de comparación de medias. La elección del tipo de test a aplicar depende de si la variable dependiente frente a la variable demográfica supera el test de normalidad. Si supera el test de normalidad se aplica el test de la T de Student, en el caso de no superar el test de normalidad, se aplica el test de la U de Mann-Whitney.

A continuación, a modo de ejemplo, se analizan dos pares de variables. En primera instancia se toma la variable dependiente *Calidad*, correspondiente al tratamiento TDD1, mientras que como variable demográfica se toma la *experiencia de programación*. La variable demográfica *experiencia de programación* en el primer caso presenta normalidad, este caso se describe en el apartado 5.1.1 (Análisis dicotómico cuando hay normalidad),

mientras que en el segundo caso no presenta normalidad, este caso se describe en el apartado 5.1.2 (Análisis dicotómico cuando no hay normalidad).

5.1.1. Análisis dicotómico cuando hay normalidad

A continuación se analiza la variable demográfica *experiencia de programación* frente a la variable dependiente *Calidad*. Lo primero que se hace es la prueba de normalidad utilizando los test de Kolmogorov-Smirnov y Shapiro-Wilk. Utilizamos la información que proporciona el test de Shapiro-Wilk en detrimento de Kolmogorov-Smirnov, puesto que se trata de un test más potente cuando la muestra tiene un tamaño menor que 50 [5].

Los resultados del test de Kolmogorov-Smirnov son p-valor = . y 0,2 para la categoría Negativa y Positiva respectivamente, tal y como se muestra en la tabla 5.1. No se proporciona un valor para la categoría negativa, esto sucede porque el tamaño de la muestra es muy pequeño como para que este test proporcione resultados.

Los resultados del de Shapiro-Wilk son p-valor = 0,084 y 0,836 para la categoría Positiva y Negativa respectivamente, tal y como se observa en la tabla 5.1. En este caso con un nivel de significación de 0,05 se puede concluir que existe normalidad.

En general, las pruebas de normalidad no resultan muy robustas cuando el tamaño muestral es muy pequeño. El nivel de significación para el caso de la categoría negativa (0,084) es muy cercano a 0,05, por lo tanto aunque en principio se acepta la hipótesis nula de que hay normalidad, se debe tener en cuenta que el nivel de significación con el que se acepta la hipótesis nula es bastante débil.

Programacion		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,305	4	.	,789	4	,084
	Positiva	,124	20	,200 [*]	,974	20	,836

Tabla 5.1. Prueba de normalidad: experiencia de programación vs Calidad

Dado que la variable Experiencia de programación presenta normalidad el test que se aplica es la T de Student para muestras independientes.

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia
Quality	Se han asumido varianzas iguales	67,570	,000	-3,626	22	,001	-36,62600	10,10223	-57,57674
	No se han asumido varianzas iguales			-1,735	3,080	,179	-36,62600	21,11047	-102,83690

Tabla 5.2. Resultado de la prueba de la T de Student de la experiencia de programación y la Calidad

La tabla 5.2 muestra el resultado de la prueba de Levene y la comparación de medias para muestras independientes al que se le aplica el test de la T de Student. El resultado de la prueba de Levene muestra que no hay igualdad de varianzas, por lo tanto nos fijamos en la segunda fila de la tabla, es decir, no se asumen varianzas iguales. El nivel de significación de el test de la T de Student es de $0,179 > 0,05$, esto es, no hay evidencia para afirmar que el grupo de sujetos con experiencia genere código con un nivel de calidad distinto que el de los sujetos con poca experiencia.

A continuación en la figura 2 se muestra un diagrama de cajas donde se representa las categorías Negativa y Positiva de la variable demográfica *experiencia de programación* frente a la variable dependiente *Calidad*, con el objetivo de tener más evidencias sobre los resultados que se acaban de obtener con el test de de la T de Student para muestras independientes.

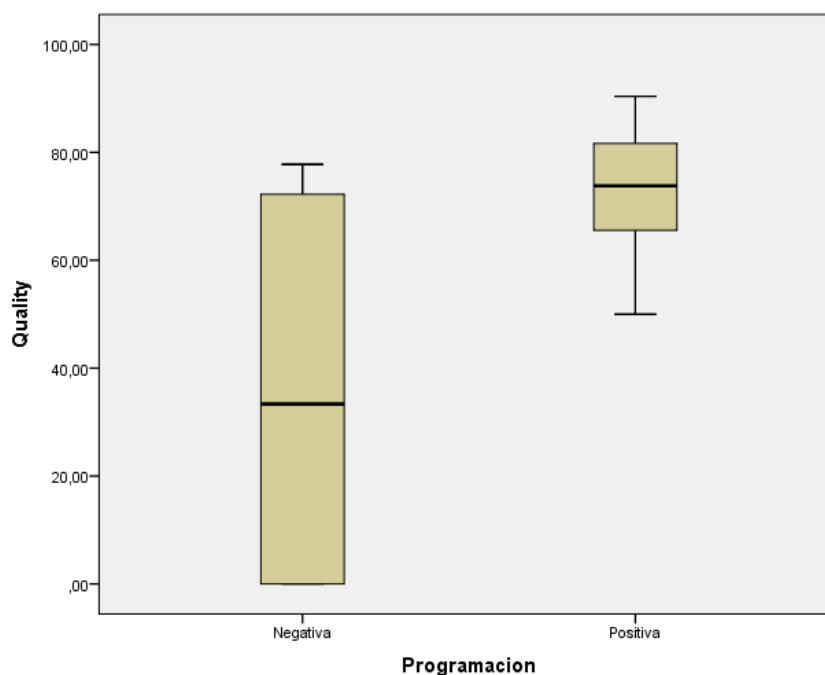


Figura 2. Diagrama de cajas: Programación vs Calidad

El test de la T de Student para muestras independientes indica que no hay evidencia para afirmar que el grupo de sujetos con experiencia en programación genere código con un nivel de calidad distinto que el de los sujetos con poca experiencia. Sin embargo el diagrama de cajas de la Figura 2 sugiere de forma bastante clara una influencia de la *experiencia de programación* en la variable respuesta *Calidad*.

Al analizar la gráfica se puede apreciar también que las varianzas del grupo Negativo son muy altas, oscila entre 0 y 50. Esto disminuye muy considerablemente el poder del test de la T de Student. Además no se observan outliers, por lo que no hay nada que se pueda

hacer para aumentar el poder del test. Por lo tanto tanto el p-valor como el diagrama de cajas señalan que la *experiencia de programación* si podría demostrarse influyente sobre la *Calidad*.

5.1.2. Análisis dicotómico cuando NO hay normalidad

A continuación se analiza la influencia que tiene la variable demográfica *experiencia académica* sobre la variable dependiente *Calidad*, correspondiente al tratamiento TDD1. Lo primero que se hace es la prueba de normalidad utilizando los test de Kolmogorov-Smirnov y Shapiro-Wilk. Utilizamos la información que proporciona el test de Shapiro-Wilk en detrimento de Kolmogorov-Smirnov, puesto que se trata de un test más potente cuando la muestra tiene una tamaño menor que 50 [5].

	Academica	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,307	4	.	,729	4	,024
	Positiva	,153	20	,200 [*]	,964	20	,618

Tabla 5.3. Prueba de normalidad: experiencia académica vs Calidad

Los resultados del test de Kolmogorov-Smirnov son p-valor = . y 0,2 para la categoría Negativa y Positiva respectivamente, tal y como se muestra en la tabla 5.3. No se proporciona un valor para la categoría negativa, esto sucede porque el tamaño de la muestra es muy pequeño como para que este test proporcione resultados.

Los resultados del de Shapiro-Wilk son p-valor = 0,024 y 0,618 para las categorías Negativa y Positiva respectivamente, como se puede apreciar en la tabla 5.3. La categoría Negativa de la variable demográfica *experiencia académica* no supera el test de normalidad (p-valor $0,024 < 0,05$), por lo tanto para la comparación de medias se utiliza el test de la U deMann-Whitney.

El resultado del test de la U de Mann-Whitney para muestras independientes es el que se muestra en la Tabla 5.4.

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de Academica.	Prueba U de Mann-Whitney de muestras independientes	,023 ¹	Rechazar la hipótesis nula.

Tabla 5.4. Resumen de la prueba de hipótesis: U de Mann-Whitney

Este test realiza un contraste de hipótesis en el que la hipótesis nula H_0 sostiene que las medianas son las mismas en las categorías Negativa y Positiva de la variable demográfica *experiencia académica*. El p-valor = $0,023 < 0,05$ lo que indica que las

diferencias en las medias es significativa, por lo tanto hay evidencias que permiten rechazar la hipótesis nula H_0 .

A continuación en la Figura 3 se muestra un diagrama de cajas donde se representa las categorías Negativa y Positiva de la variable demográfica *experiencia académica* frente a la variable dependiente *Calidad*, con el objetivo de tener más evidencias sobre los resultados que se han obtenido con el test de la U de Mann-Whitney para muestras independientes.

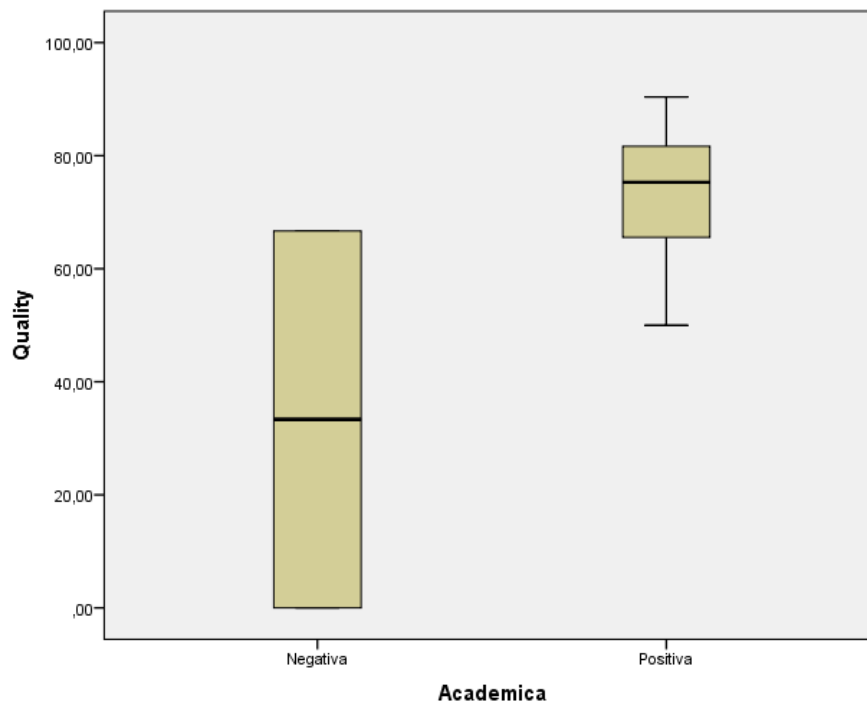


Figura 3. Diagrama de cajas: experiencia académica vs Calidad

El diagrama de cajas de la Figura 3 muestra de forma visible una clara influencia de la *Experiencia académica* en la variable respuesta *Calidad* correspondiente a los experimentos realizados por los sujetos de TDD1, tal y como confirma test de la U de Mann-Whitney.

5.1.3. Resumen de resultados

El proceso anterior se ha realizado para los 72 pares de variables que componen el análisis dicotómico. El detalle completo de este análisis se encuentra en el Anexo B. Análisis dicotómico. La Tabla 5.6 muestra el resumen de los resultados obtenidos en este análisis. Muestra los p-valores del resultado del test de comparación de medias correspondiente (T de Student o U de Mann-Whitney).

Los valores sombreados se corresponden con el test de la T de Student, el resto se corresponden con el test de la U de Mann-Whitney. Los valores que aparecen en negrita son los que resultan significativos, es decir, los que según el test aplicado, denotan que

existe influencia de la variable demográfica sobre la variable dependiente. A su vez, los p-valores en negrita se han coloreado en rojo o verde para denotar significación estadística al 5% y 10%, respectivamente.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	0,135	0,002	0,451	0,183	0,023	0,446	0,097	0,003	0,898
Programación	0,241	0,005	0,744	0,347	0,179	0,852	0,241	0,003	0,562
Java	0,241	0,241	0,388	0,273	0,737	0,932	0,183	0,183	0,514
Testing	0,009	0,038	0,908	0,052	0,045	0,649	0,007	0,61	0,535
JUnit	0,599	0,318	0,5	0,815	0,726	0,496	0,379	0,29	0,934
IDE	1	0,068	0,863	0,627	0,241	0,143	0,911	0,135	0,907
MetDes	0,581	0,103	0,051	0,945	0,16	0,552	0,679	0,12	0,185
TDD	0,89	0,41	0,614	0,178	0,045	0,257	0,16	0,63	0,566

Tabla 5.6. Resultados de la comparación de medias: T de Student y U de Mann-Whitney

El tratamiento TDD1 es el que presenta más diferencias significativas en el test de comparación de medias. Se han encontrado diferencias en 11 de los 24 pares de variables que se han comparado. El tratamiento TLD presenta menos diferencias, en concreto 4 de los 24 pares de variables que se han comparado. Por su parte el tratamiento TDD2 no presenta ninguna diferencia. Estos resultados indican que las variables demográficas (ej. Experiencia *académica*, en *TDD*, etc.) influyen más en los resultados obtenidos en el experimento cuando el tratamiento es TDD1. También se puede decir a la vista de los resultados que para el tratamiento TDD2 las variables demográficas no le afectan en ningún sentido.

También se puede destacar el hecho de que las variables dependientes *Productividad* y *Calidad* del tratamiento TDD2 superan el test de normalidad para todas las variables demográficas.

A la vista de los resultados se parece que las variables demográficas que más influencia provocan en las variables independientes son la *experiencia académica* y la experiencia en *Testing*. Del mismo modo se puede decir que las variables demográficas experiencia en *Java* y en *JUnit* no afectan a ninguna de las variables dependientes.

La Tabla 6.2 muestra el tamaño de efecto de la comparación de medias.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	-0.402	-0.348	-0.787	-0.223	-0.333	-0.524	-0.949	-0.943	-0.978
Programación	-0.406	-0.346	-0.739	-0.219	-0.310	-0.456	-0.948	-0.943	-0.973
Java	-0.379	-0.249	-0.790	-0.213	-0.300	-0.574	-0.941	-0.898	-0.963
Testing	-0.245	-0.130	-0.552	-0.124	-0.163	-0.269	-0.861	-0.713	-0.939
JUnit	-0.152	-0.094	-0.456	-0.084	-0.124	-0.250	-0.722	-0.593	-0.879
IDE	-0.388	-0.328	-0.604	-0.233	-0.394	-0.430	-0.904	-0.920	-0.968
MetDes	-0.295	-0.218	-0.757	-0.183	-0.257	-0.477	-0.886	-0.864	-0.975
TDD	0	0	-0.068	0	0	-0.034	0	0	-0.268

Tabla 5.7 Tamaño de efecto de la comparación de medias

El nivel de significación nos determina si la diferencia de medias es lo suficientemente considerable como para afirmar que no son iguales a un cierto nivel de significación. Por otro lado, el tamaño de efecto determina de forma cuantitativa el efecto de esa diferencia encontrada.

5.2. Análisis de correlación de Pearson

El segundo análisis que se ha llevado a cabo es la correlación de Pearson. Este análisis se ha realizado con los datos cuantitativos obtenidos en el análisis demográfico, es decir, sin dicotomizar.

La razón para realizar este análisis viene determinada porque la dicotomización representa en muchas ocasiones una partición arbitraria de un conjunto de datos, lo que ocasiona que las inferencias realizadas sobre datos dicotomizados sean erróneas [3]. La correlación de Pearson, al realizarse sobre los datos originales, no sufre este problema (aunque posee otros problemas que descubriremos posteriormente).

El número de tests que es necesario realizar en el análisis de correlación es similar al análisis dicotómico (72 tests). Por este motivo procedemos al igual que en la sección anterior: mostramos dos ejemplos de análisis y posteriormente mostramos una tabla de resumen global.

A continuación, a modo de ejemplo se realiza un estudio de dos casos, uno para cuando las variables están relacionadas y otro en el que no se manifiesta ninguna relación. Igual que en el análisis dicotómico se analiza una variable demográfica (ej. *experiencia académica*, en *TDD*, etc.) frente a una variable dependiente (ej. *Productividad*, *Calidad*, etc.). En este caso la variable dependiente *Calidad* se toma del tratamiento TLD.

5.2.1. Variables correlacionadas

La Tabla 5.8 muestra el resultado de la correlación de Pearson para la variable demográfica *experiencia en TDD* y la variable dependiente *Calidad*.

		Quality	TDD
Quality	Correlación de Pearson	1	,409 [*]
	Sig. (bilateral)		,047
	N	24	24
TDD	Correlación de Pearson	,409 [*]	1
	Sig. (bilateral)	,047	
	N	24	24

Tabla 5.8. Correlación: experiencia en TDD vs Calidad

Se observa en la correlación entre la *Calidad* y la *experiencia en TDD*, que el p-valor del coeficiente de correlación de Pearson que se obtiene es de 0,047, lo que implica que es

estadísticamente significativo al nivel del 5%. Ello implica que las variables *experiencia en TDD* y *Calidad* están correlacionadas.

El coeficiente de correlación de Pearson es de 0,409. Estamos frente a un caso de correlación media según Cohem.

La Figura 4 muestra el gráfico de dispersión para las variables *Calidad* y *experiencia en TDD*. El gráfico de dispersión permite apreciar que hay cierta relación entre las variables *Calidad* y *experiencia en TDD*, lo cual refuerza las conclusiones del test de correlación de Pearson. Sin embargo aunque el resultado afirma que las variables están correlacionadas, la bondad del ajuste lineal es muy bajo ($R^2 = 0,167$), solo tiene un poder explicativo del 16,7%. El motivo por el que se obtienen estos resultados es que la variable demográfica *experiencia en TDD* por sí sola no llega a predecir correctamente la *Calidad*, dado que hay más variables involucradas, lo que sugiere la necesidad de pasar a un análisis de correlación a la regresión múltiple.

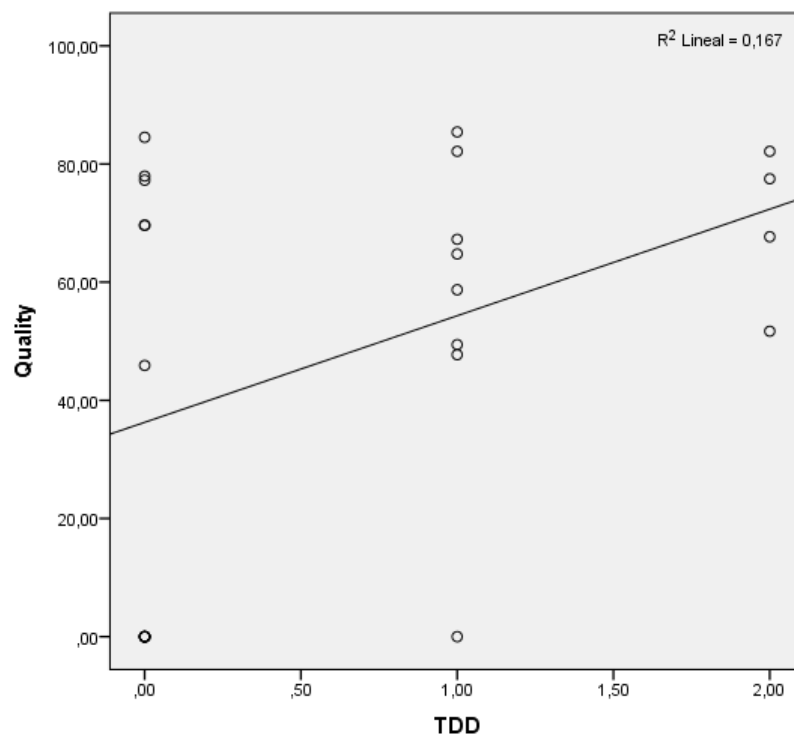


Figura 4. Gráfico de dispersión: experiencia en TDD vs Calidad

5.2.2. Variables NO correlacionadas

La figura 5.9 muestra el resultado de la correlación de Pearson para la variable demográfica *experiencia académica* y la variable dependiente *Calidad*.

		Quality	Academica
Quality	Correlación de Pearson	1	-,086
	Sig. (bilateral)		,690
	N	24	24
Academica	Correlación de Pearson	-,086	1
	Sig. (bilateral)	,690	
	N	24	24

Tabla 5.9. Correlación: experiencia académica vs Calidad

Se observa en la correlación entre la *Calidad* y la *experiencia académica*, que el p-valor que se obtiene es de 0,690, que es mayor que 0,05, nivel de significación usado generalmente por los investigadores. Es incluso mayor que 0,1, que un es nivel de significación conservador que se ha decidido considerar para realizar este análisis Post-Hoc. Por tanto no podemos afirmar que haya correlación entre *Calidad* y *experiencia académica*, es decir se acepta la hipótesis nula de que no existe correlación entre las variables.

Sobre el coeficiente de correlación de Pearson podemos decir que es muy bajo (-0,086). En el caso de ser significativo estaríamos frente a una correlación baja, según Cohen. Además el signo negativo del coeficiente muestra que la relación es inversa, es decir un incremento en *experiencia académica* provocaría un decremento en la *Calidad*. Lo cual no parece muy razonable.

La Figura 5 muestra el gráfico de dispersión para las variables *Calidad* y *experiencia académica*.

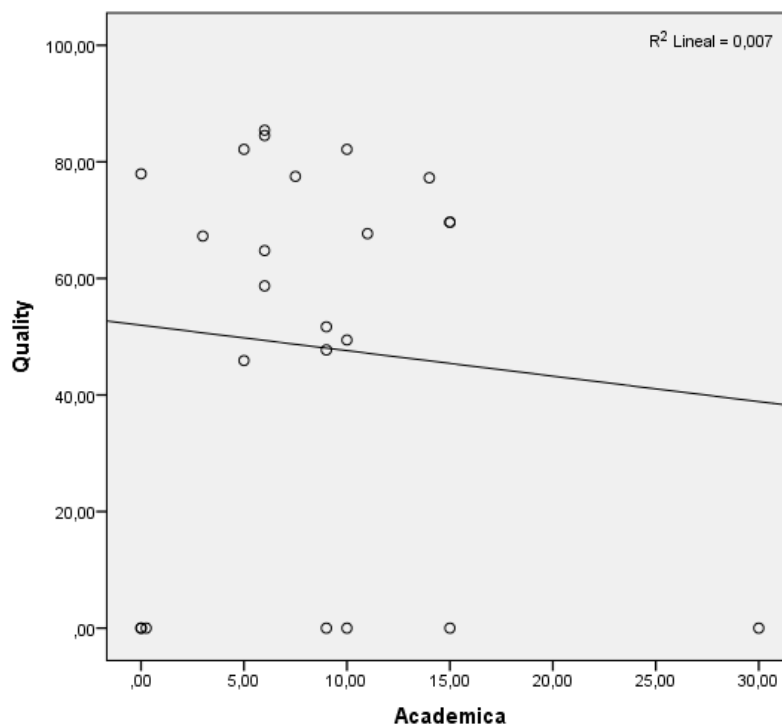


Figura 5. Gráfico de dispersión: experiencia académica vs Calidad

En el gráfico de dispersión se puede apreciar que no hay una relación evidente entre las variables, lo que refuerza el resultado obtenido en el test de correlación. La bondad de ajuste es prácticamente cero (0,007), lo que resulta perfectamente compatible con la falta de correlación observada entre la variable demográfica y la variable dependiente.

5.2.3. Resumen de resultados

El procedimiento anterior se ha realizado con todos los pares de variables para cada uno de los tratamientos. A continuación se presenta un resumen de los resultados obtenidos. El análisis completo se encuentra en el Anexo C. Análisis de correlación de Pearson.

La Tabla 5.10 muestra los p-valores o significación de cada uno de los pares de variables que se han analizado. Los valores que aparecen en negrita son los que resultan significativos, es decir, los que según el test aplicado, denotan que existe influencia de la variable demográfica sobre la variable dependiente. A su vez, los p-valores en negrita se han coloreado en rojo o verde para denotar significación estadística al 5% y 10%, respectivamente.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	,413	,706	,168	,690	,232	,753	,479	,179	,450
Programación	,789	,280	,251	,650	,140	,480	,848	,054	,443
Java	,793	,333	,573	,438	,936	,817	,703	,040	,511
Testing	,557	,163	,050	,601	,608	,483	,621	,020	,364
JUnit	,146	,167	,710	,068	,616	,771	,104	,020	,577
IDE	,777	,100	,769	,969	,284	,143	,600	,128	,804
MetDes	,064	,306	,051	,589	,102	,552	,139	,166	,243
TDD	,153	,306	,924	,047	,084	,133	,099	,398	,681

Tabla 5.10. p-valores de la correlación de Pearson

A continuación, en Tabla 5.11 se muestra el coeficiente de correlación de Pearson para todos los pares de variables que se han analizado. El código de colores es similar al utilizado en la Tabla 5.10.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	-,175	,081	-,297	-,086	,253	,069	-,15	,283	-,162
Programación	-,058	,230	-,249	,098	,311	-,155	-,04	,398	-,168
Java	,057	,206	-,124	,166	,017	,051	,082	,423	-,145
Testing	,126	,294	-,414	,112	,110	-,154	,106	,471	-,198
JUnit	,306	,292	-,082	,379	,108	-,064	,340	,471	-,123
IDE	-,061	-,344	,065	,008	-,228	,315	-,11	-,32	-,055
MetDes	-,384	,218	-,411	-,116	,342	,131	-,31	,292	-,253
TDD	,301	,218	-,021	,409	,360	,323	,344	,181	-,090

Tabla 5.11. Coeficientes de correlación de Pearson

Ninguno de los análisis ha arrojado un coeficiente de correlación superior a 0,5, es decir, ningún par de variables presenta una correlación alta (los valores se encuentran en el intervalo 0,2 – 0,5 que según Cohen corresponde a una correlación media). Resulta curioso que haya valores negativos en el coeficiente de correlación de Pearson, eso significa que tener experiencia en la variable demográfica influye en sentido inverso al que por sentido común sería esperable en la variable respuesta correspondiente. Por ejemplo según los resultados que muestra la Tabla 5.11, tener experiencia en *Testing* influye negativamente en la *Productividad* para el tratamiento de TDD2. Desde luego, cuando menos hay que dudar de este resultado, que puede deberse al pequeño tamaño de la muestra o simple casualidad.

Igual que en el análisis dicotómico el tratamiento TDD1 es el que presenta mayor número de pares de variables correlacionadas. En concreto tiene 7 pares de variables que están relacionadas. Por su parte el tratamiento TLD tiene 5 pares de variables que presentan correlación. Finalmente TDD2 presenta dos pares de variables correlacionadas.

Las variables demográficas que más influyen en las variables dependientes por su reiteración en la aparición en los diferentes tratamientos, son: la *experiencia JUnit*, *metodologías de desarrollo* y *experiencia en TDD*.

5.3. Análisis de regresión lineal múltiple

Los métodos analizados en los apartados anteriores no son del todo apropiados para el tipo de análisis que nos ocupa. Esto se debe a que tanto el análisis dicotómico como el análisis de correlación de Pearson, analizan cada variable demográfica por separado. Cuando varias variables demográficas pueden actuar simultáneamente, puede dar lugar a falsos positivos y falsos negativos.

Por lo tanto se ha llevado a cabo un tercer tipo de análisis. Se trata del análisis de regresión lineal múltiple. El análisis de regresión permite considerar todas las variables simultáneamente, evitando los efectos perniciosos de las comparaciones por pares [6]. En contrapartida tenemos el problema de que la muestra es muy pequeña para que el análisis de regresión sea fiable [7].

Para realizar este análisis se han introducido en el programa estadístico todas las variables demográficas por cada una de las variables respuesta del experimento, lo que reduce considerablemente el número de tests que se han realizado.

Si con la dicotomización o la correlación se han realizado 72 tests, con la regresión se han realizado nueve, uno por cada variable respuesta de los tres tratamientos. El resumen de resultados se muestra en la Tabla 5.12.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	,776	,542	,413	,881	,097	,289	,602	,795	,736
Programación	,616	,867	,368	,640	,265	,111	,895	,815	,775
Java	,281	,781	,745	,566	,058	,447	,353	,707	,929
Testing	,814	,416	,033	,713	,820	,140	,788	,463	,574
Junit	,181	,795	,149	,186	,467	,569	,115	,901	,722
IDE	,660	,196	,952	,560	,330	,654	,969	,177	,799
MetDes	,034	,428	,201	,291	,928	,539	,078	,450	,510
TDD	,158	,536	,876	,123	,042	,151	,150	,701	,826

Tabla 5.12. p-valores de la regresión lineal múltiple

Los valores de la Tabla 5.12 se corresponden con el p-valor de la estimación de las betas correspondientes a cada variable demográfica. Como se ha comentado antes, este es el resultado del análisis de regresión lineal introduciendo todas las variables en el modelo (entrada forzosa). Los valores que aparecen en negrita son los que resultan significativos, es decir, los que según el test aplicado, denotan que existe influencia de la variable demográfica sobre la variable dependiente. A su vez, los p-valores en negrita se han coloreado en rojo o verde para denotar significación estadística al 5% y 10%, respectivamente.

A la vista de los resultados de la Tabla 5.12, hay pocas variables demográficas que resultan significativas en el modelo de regresión lineal. En el tratamiento de TLD la experiencia en *metodologías de desarrollo* influye en la *Productividad* y la *Calidad*. En el tratamiento TDD1 la variable respuesta *Calidad* se ve influida por la experiencia *académica*, en *Java* y en *TDD*. En el tratamiento TDD2 la variable *Productividad* se ve influida por la variable experiencia en *Testing*.

En la Tabla 5.13 se muestra la estimación de las betas correspondientes a cada variable demográfica frente a cada una de las variables dependientes. El código de colores es similar al utilizado en la Tabla 5.12.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	,323	-1,308	,429	,326	2,28	1,165	,093	-,067	,054
Programación	,932	,584	-,828	1,671	2,448	-3,152	,039	,098	-,080
Java	-2,809	1,340	,327	-2,837	-5,99	1,604	-,379	,217	-,028
Testing	,617	4,055	-2,45	-1,854	-,690	-3,387	-,111	,438	-,185
Junit	4,171	-1,483	1,868	7,947	2,576	1,481	,783	,085	,137
IDE	4,622	-26,312	,262	11,816	-12,03	4,095	,064	-3,296	-,345
MetDes	-28,78	18,900	-6,514	-25,939	1,306	6,359	-3,67	2,156	-1,014
TDD	7,549	6,071	-,324	15,990	13,07	6,428	1,213	,449	-,141

Tabla 5.13. Betas de la regresión lineal múltiple

Los valores de las betas que se pueden observar en la Tabla 5.13 sorprenden. La mayoría de las variables demográficas que presentan un p-valor significativo les corresponde un valor de la beta del modelo de regresión que resulta ser negativa. La interpretación de los resultados sería la siguiente: la variable demográfica *experiencia*

académica influye positivamente en la *Calidad* del tratamiento TDD1, y el valor de la beta es 2,28. La variable demográfica *experiencia en Testing* influye negativamente en la *Productividad* del tratamiento TDD2, y el valor de la beta es -2,45. Así con todas las variables que resultan nivel de significación menor que 0.1.

5.4. Stepwise regression (paso a paso hacia atrás)

Para afinar más los resultados que se obtienen, se ha realizado un análisis de regresión lineal múltiple paso a paso (stepwise regression). Consiste en ir introduciendo todas las variables y realizando iteraciones en las que se va quitando del modelo la variable demográfica que presenta menor capacidad de predicción.

La stepwise regression es un método de reducción de dimensiones, que al quitar variables en el modelo produce un aumento del número de sujetos/variable, resultando de esta forma más fiable el análisis [7].

Este apartado consta de tres partes, una por cada variable respuesta del experimento (*Productividad*, *Calidad*, *TUS*). En cada apartado se presentan los modelos generados mediante la stepwise regression para cada uno de los tratamientos TLD, TDD1 y TDD2.

5.4.1. Influencia de las variables demográficas sobre la *Productividad*

En este apartado se realiza un análisis de la influencia de las variables demográficas sobre la variable respuesta *Productividad*. Se muestra el modelo de regresión elegido en cada caso para cada uno de los tres tratamientos: TLD, TDD1 y TDD2.

5.4.1.1. Tratamiento TLD

La Tabla 5.14 muestra los coeficientes (betas) de los modelos de regresión obtenidos para la variable respuesta *Productividad* correspondientes al tratamiento TLD. Muestra también las variables que componen cada modelo, así como el nivel de significación que tiene cada variable en el modelo correspondiente. Las variables demográficas van mostrando tendencia a la significación a medida que avanza la stepwise regression son: la *Experiencia académica*, la *Experiencia en JUnit* y la *Experiencia en metodologías de desarrollo*, y la *Experiencia en TDD*.

La Tabla 5.15 muestra el resumen de los modelos de regresión que en cada paso va arrojando la stepwise regression. En este caso se trata de siete modelos diferentes, que en la tabla aparecen numerados del 1 al 7.

Para elegir el modelo se tiene en cuenta, por un lado, el coeficiente de determinación R^2 , que es el que representa el poder explicativo del modelo. Por otro lado también se ha tenido en cuenta el número de variables predictivas que resultan significativas. Dado que la hipótesis que se intenta demostrar es que las variables demográficas influyen en la variable

dependiente, nos interesa tener el mayor número de variables significativas en el modelo. Atendiendo a estos criterios se elige el modelo 6, que aparece sombreado en la Tabla 5.14.

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
1 (Constante)	25,400	14,145		1,796	,093
Academica	,323	1,113	,114	,290	,776
Programacion	,932	1,821	,156	,512	,616
Java	-2,809	2,513	-,485	-1,118	,281
Testing	,617	2,572	,089	,240	,814
JUnit	4,171	2,976	,545	1,402	,181
IDE	4,622	10,300	,094	,449	,660
MetDes	-28,784	12,304	-,587	-2,339	,034
TDD	7,549	5,084	,308	1,485	,158
2 (Constante)	26,249	13,286		1,976	,066
Academica	,463	,920	,163	,503	,622
Programacion	,802	1,688	,135	,475	,641
Java	-2,792	2,437	-,482	-1,146	,269
JUnit	4,608	2,285	,602	2,016	,061
IDE	4,290	9,901	,087	,433	,671
MetDes	-28,771	11,936	-,586	-2,411	,028
TDD	7,426	4,907	,303	1,513	,150
3 (Constante)	30,014	9,809		3,060	,007
Academica	,491	,895	,173	,549	,590
Programacion	,597	1,580	,100	,377	,711
Java	-2,632	2,350	-,454	-1,120	,278
JUnit	4,474	2,210	,584	2,025	,059
MetDes	-27,877	11,472	-,568	-2,430	,026
TDD	7,308	4,781	,298	1,529	,145
4 (Constante)	31,718	8,498		3,732	,002
Academica	,503	,873	,178	,577	,571
Java	-2,337	2,164	-,403	-1,080	,294
JUnit	4,524	2,152	,591	2,102	,050
MetDes	-26,337	10,463	-,537	-2,517	,022
TDD	7,090	4,632	,289	1,531	,143
5 (Constante)	31,912	8,341		3,826	,001
Java	-1,416	1,433	-,244	-,988	,336
JUnit	4,009	1,923	,524	2,084	,051
MetDes	-23,351	8,930	-,476	-2,615	,017
TDD	6,531	4,449	,266	1,468	,158
6 (Constante)	30,387	8,192		3,709	,001
JUnit	2,724	1,417	,356	1,923	,069
MetDes	-24,125	8,890	-,492	-2,714	,013
TDD	6,616	4,445	,270	1,488	,152
7 (Constante)	33,474	8,151		4,107	,001
JUnit	3,167	1,425	,414	2,223	,037
MetDes	-23,444	9,132	-,478	-2,567	,018

Tabla 5.14. Pasos de la stepwise regression para la productividad en el tratamiento TLD

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,658 ^a	,434	,131	17,41047
2	,657 ^b	,431	,183	16,88995
3	,652 ^c	,425	,222	16,48153
4	,648 ^d	,420	,259	16,08414
5	,640 ^e	,409	,285	15,79906
6	,615 ^f	,379	,286	15,78966
7	,557 ^g	,310	,244	16,24008

Tabla 5.15. Resumen de los modelos de regresión para la productividad en el tratamiento TLD

Las variables demográficas que son significativas para el modelo 6 son: la *experiencia en JUnit* y la *experiencia en metodologías de desarrollo*, con un p-valor de 0,069 y 0,013 respectivamente.

5.4.1.2. Tratamiento TDD1

La Tabla 5.16 muestra los coeficientes (betas) de los modelos de regresión obtenidos para la variable respuesta *Productividad* correspondiente al tratamiento TDD1. Las variables demográficas que van mostrando significación a medida que avanza la stepwise regression son: la *Experiencia en Testing*, la *Experiencia en IDE* y la *Experiencia en TDD*.

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
1 (Constante)	33,951	26,675		1,273	,222
Academica	-1,308	2,098	-,276	-,623	,542
Programacion	,584	3,435	,059	,170	,867
Java	1,340	4,739	,138	,283	,781
Testing	4,055	4,851	,348	,836	,416
JUnit	-1,483	5,612	-,116	-,264	,795
IDE	-26,312	19,423	-,320	-1,355	,196
MetDes	18,900	23,203	,230	,815	,428
TDD	6,071	9,588	,148	,633	,536
2 (Constante)	36,461	21,534		1,693	,110
Academica	-1,237	1,993	-,261	-,621	,544
Java	1,622	4,303	,167	,377	,711
Testing	3,811	4,491	,327	,849	,409
JUnit	-1,297	5,335	-,101	-,243	,811
IDE	-27,324	17,920	-,333	-1,525	,147
MetDes	20,355	20,903	,248	,974	,345
TDD	5,819	9,181	,142	,634	,535
3 (Constante)	36,670	20,913		1,753	,098
Academica	-,959	1,586	-,202	-,605	,553
Java	,963	3,249	,099	,296	,771
Testing	3,166	3,523	,272	,899	,381
IDE	-26,933	17,347	-,328	-1,553	,139
MetDes	18,830	19,379	,229	,972	,345
TDD	5,636	8,893	,137	,634	,535
4 (Constante)	36,854	20,367		1,809	,087
Academica	-,726	1,343	-,153	-,541	,595
Testing	3,644	3,052	,313	,194	,248
IDE	-26,956	16,902	-,328	-1,595	,128
MetDes	17,786	18,567	,217	,958	,351
TDD	6,278	8,404	,153	,747	,465
5 (Constante)	37,220	19,974		1,863	,078
Testing	2,654	2,395	,228	,108	,282
IDE	-28,561	16,326	-,348	-1,749	,096
MetDes	14,216	17,028	,173	,835	,414
TDD	6,939	8,158	,169	,851	,406
6 (Constante)	45,896	16,927		2,711	,013
Testing	3,205	2,285	,275	,1403	,176
IDE	-26,994	16,095	-,329	-1,677	,109
TDD	7,565	8,062	,184	,938	,359
7 (Constante)	51,199	15,910		3,218	,004
Testing	3,340	2,274	,287	,1469	,157
IDE	-27,730	16,030	-,338	-1,730	,098
8 (Constante)	60,385	15,010		4,023	,001
IDE	-28,232	16,442	-,344	-1,717	,100
9 (Constante)	36,859	6,382		5,775	,000

Tabla 5.16. Pasos de la stepwise regression para la productividad en el tratamiento TDD1

La Tabla 5.17 muestra el resumen de los modelos de regresión que en cada paso va arrojando la stepwise regression para la variable respuesta *Productividad*. En este caso son 9 modelos. El modelo elegido en este caso es el 7 siguiendo el mismo criterio que en el apartado 5.4.1.1.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,530 ^a	,281	-,103	32,83345
2	,529 ^b	,279	-,036	31,82149
3	,526 ^c	,277	,021	30,92833
4	,522 ^d	,273	,071	30,13450
5	,511 ^e	,261	,106	29,56808
6	,484 ^f	,234	,119	29,34324
7	,448 ^g	,200	,124	29,25963
8	,344 ^h	,118	,078	30,01951
9	,000 ⁱ	,000	0,000	31,26499

Tabla 5.17. Resumen de los modelos de regresión para la productividad en el tratamiento TDD1

En el modelo que se ha elegido (7) la variable demográfica que es significativa es la *experiencia en IDE*.

5.4.1.3. Tratamiento TDD2

La Tabla 5.18 muestra los coeficientes (betas) de los modelos de regresión obtenidos para la variable respuesta *Productividad* correspondientes al tratamiento TDD2. Las variables demográficas que van mostrando significación a medida que avanza la stepwise regression son: la *Experiencia académica*, la *Experiencia en Testing*, la *Experiencia en JUnit* y la *Experiencia en Metodologías de Desarrollo*.

La Tabla 5.19 muestra el resumen de los modelos de regresión que en cada paso va arrojando la stepwise regression para la variable respuesta *Productividad*. En este caso son 8 modelos. El modelo elegido en este caso es el 4, la elección del modelo se ha realizado siguiendo el mismo criterio que en el apartado 5.4.1.1.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,655 ^a	,429	,103	6,81649
2	,655 ^b	,429	,162	6,58624
3	,654 ^c	,428	,213	6,38285
4	,651 ^d	,423	,254	6,21550
5	,618 ^e	,382	,244	6,25542
6	,594 ^f	,353	,251	6,22742
7	,511 ^g	,261	,187	6,48759
8	,414 ^h	,171	,132	6,70435

Tabla 5.19. Resumen de los modelos de regresión para la productividad en el tratamiento TDD2

En este modelo las variables significativas son la *experiencia en Testing*, *experiencia en JUnit* y la *experiencia en metodologías de desarrollo*.

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
1 (Constante)	28,752	5,969		4,817	,000
Academica	,429	,508	,386	,844	,413
Programacion	-,828	,891	-,359	-,930	,368
Java	,327	,986	,147	,332	,745
Testing	-2,456	1,042	-,916	-2,358	,033
JUnit	1,868	1,222	,640	1,529	,149
IDE	,262	4,295	,014	,061	,952
MetDes	-6,514	4,850	-,351	-1,343	,201
TDD	-,324	2,034	-,034	-,160	,876
2 (Constante)	29,009	4,095		7,085	,000
Academica	,437	,471	,394	,929	,368
Programacion	-,852	,774	-,369	-1,100	,289
Java	,337	,939	,151	,359	,725
Testing	-2,469	,984	-,921	-2,509	,024
JUnit	1,874	1,177	,642	1,593	,132
MetDes	-6,478	4,652	-,349	-1,393	,184
TDD	-,323	1,965	-,034	-,165	,871
3 (Constante)	28,849	3,854		7,485	,000
Academica	,454	,446	,409	1,016	,325
Programacion	-,851	,750	-,369	-1,135	,273
Java	,311	,897	,140	,347	,733
Testing	-2,462	,953	-,918	-2,584	,020
JUnit	1,866	1,139	,640	1,638	,121
MetDes	-6,638	4,410	-,357	-1,505	,152
4 (Constante)	28,805	3,751		7,679	,000
Academica	,531	,377	,478	1,409	,177
Programacion	-,779	,702	-,337	-1,110	,282
Testing	-2,462	,928	-,918	-2,654	,017
JUnit	2,065	,958	,708	2,156	,046
MetDes	-7,199	3,995	-,388	-1,802	,089
5 (Constante)	26,561	3,180		8,352	,000
Academica	,270	,296	,243	,911	,374
Testing	-2,108	,877	-,786	-2,405	,027
JUnit	1,467	,796	,503	1,841	,082
MetDes	-7,664	3,998	-,413	-1,917	,071
6 (Constante)	26,760	3,158		8,473	,000
Testing	-1,636	,704	-,610	-2,324	,031
JUnit	1,238	,752	,424	1,645	,116
MetDes	-6,110	3,600	-,329	-1,697	,106
7 (Constante)	26,596	3,289		8,087	,000
Testing	-,854	,541	-,318	-1,578	,130
MetDes	-5,836	3,746	-,314	-1,558	,135
8 (Constante)	22,477	2,021		11,124	,000
Testing	-1,110	,533	-,414	-2,084	,050

Tabla 5.18. Pasos de la stepwise regression para la productividad en el tratamiento TDD2

5.4.2. Influencia de las variables demográficas sobre la *Calidad*

En este apartado se realiza un análisis de la influencia de las variables demográficas sobre la variable respuesta *Calidad*, se muestra el modelo de regresión elegido en cada paso para cada uno de los tres tratamientos: TLD, TDD1 y TDD2.

5.4.2.1. Tratamiento TLD

La Tabla 5.20 muestra los coeficientes (betas) de los modelos de regresión obtenidos para la variable respuesta *Calidad* correspondientes al tratamiento TLD. Las variables demográficas que van mostrando significación a medida que avanza la stepwise regression

son: la *Experiencia de Programación*, la *Experiencia en JUnit*, la *Experiencia en Metodologías de Desarrollo* y la *Experiencia en TDD*.

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
1 (Constante)	35,596	27,223		1,308	,211
Academica	,326	2,141	,064	,152	,881
Programacion	1,671	3,505	,156	,477	,640
Java	-2,837	4,837	-,272	-,587	,566
Testing	-1,854	4,951	-,148	-,374	,713
JUnit	7,947	5,728	,578	1,387	,186
IDE	11,816	19,823	,134	,596	,560
MetDes	-25,939	23,680	-,294	-1,095	,291
TDD	15,990	9,785	,363	1,634	,123
2 (Constante)	34,940	26,048		1,341	,199
Programacion	1,777	3,328	,166	,534	,601
Java	-2,431	3,912	-,233	-,621	,543
Testing	-1,459	4,088	-,117	-,357	,726
JUnit	7,431	4,478	,540	1,660	,116
IDE	12,211	19,043	,138	,641	,530
MetDes	-24,627	21,375	-,279	-1,152	,266
TDD	15,818	9,418	,359	1,680	,112
3 (Constante)	33,092	24,864		1,331	,201
Programacion	2,049	3,156	,191	,649	,525
Java	-3,051	3,414	-,293	-,893	,384
JUnit	6,729	3,919	,489	1,717	,104
IDE	12,740	18,492	,144	,689	,500
MetDes	-26,496	20,185	-,300	-1,313	,207
TDD	16,458	9,005	,373	1,828	,085
4 (Constante)	41,449	20,930		1,980	,063
Java	-1,881	2,854	-,181	-,659	,518
JUnit	6,721	3,855	,488	1,743	,098
IDE	9,393	17,471	,107	,538	,597
MetDes	-20,561	17,707	-,233	-1,161	,261
TDD	15,609	8,765	,354	1,781	,092
5 (Constante)	48,423	16,115		3,005	,007
Java	-1,654	2,769	-,159	-,597	,557
JUnit	6,335	3,716	,460	1,705	,105
MetDes	-19,451	17,254	-,221	-1,127	,274
TDD	15,456	8,595	,351	1,798	,088
6 (Constante)	46,642	15,580		2,994	,007
JUnit	4,835	2,694	,351	1,794	,088
MetDes	-20,355	16,909	-,231	-1,204	,243
TDD	15,556	8,454	,353	1,840	,081
7 (Constante)	31,193	8,929		3,494	,002
JUnit	4,153	2,662	,302	1,560	,134
TDD	15,032	8,533	,341	1,762	,093
8 (Constante)	36,281	8,578		4,229	,000
TDD	18,043	8,578	,409	2,103	,047

Tabla 5.20. Pasos de la stepwise regression para la calidad en el tratamiento TLD

La Tabla 5.21 muestra el resumen de los modelos de regresión que en cada paso va arrojando la stepwise regression para la variable respuesta *Calidad*. En este caso son 8 modelos. El modelo elegido en este caso es el 6, la elección del modelo se ha realizado siguiendo el mismo criterio que en el apartado 5.4.1.1.

En el modelo elegido las variables demográficas que son significativas son la *experiencia en JUnit* y la *experiencia en IDE*.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,592 ^a	,350	,004	33,50848
2	,591 ^b	,349	,065	32,46955
3	,587 ^c	,344	,113	31,62523
4	,573 ^d	,328	,141	31,11306
5	,563 ^e	,317	,173	30,52541
6	,552 ^f	,304	,200	30,03066
7	,504 ^g	,254	,183	30,35018
8	,409 ^h	,167	,130	31,32339

Tabla 5.21. Resumen de los modelos de regresión para la calidad en el tratamiento TLD

5.4.2.2. Tratamiento TDD1

La Tabla 5.22 muestra los coeficientes (betas) de los modelos de regresión obtenidos para la variable respuesta *Calidad* correspondientes al tratamiento TDD1. Las variables demográficas que van mostrando significación a medida que avanza la stepwise regression son: la *Experiencia Académica*, la *Experiencia en Programación*, la *Experiencia en Java*, la *Experiencia en IDE* y la *Experiencia en TDD*.

Modelo	Coeficientes no estandarizados		Coeficientes tipificados		t	Sig.
	B	Error tip.	Beta			
1 (Constante)	47,188	16,409			2,876	,012
Academica	2,284	1,291	,660		1,769	,097
Programacion	2,448	2,113	,336		1,159	,265
Java	-5,995	2,915	-,848		-2,056	,058
Testing	-,690	2,984	-,081		-,231	,820
JUnit	2,576	3,452	,276		,746	,467
IDE	-12,035	11,948	-,201		-1,007	,330
MetDes	1,306	14,273	,022		,092	,928
TDD	13,079	5,898	,437		2,218	,042
2 (Constante)	47,444	15,660			3,030	,008
Academica	2,327	1,165	,672		1,998	,063
Programacion	2,519	1,902	,346		1,325	,204
Java	-6,105	2,574	-,863		-2,372	,031
Testing	-,689	2,890	-,081		-,238	,815
JUnit	2,640	3,275	,283		,806	,432
IDE	-11,847	11,399	-,198		-1,039	,314
TDD	13,180	5,611	,440		2,349	,032
3 (Constante)	46,495	14,718			3,159	,006
Academica	2,170	,935	,627		2,321	,033
Programacion	2,663	1,754	,366		1,518	,147
Java	-6,122	2,501	-,866		-2,448	,026
JUnit	2,152	2,485	,230		,866	,399
IDE	-11,479	10,976	-,192		-1,046	,310
TDD	13,316	5,425	,445		2,454	,025
4 (Constante)	47,321	14,585			3,245	,004
Academica	1,900	,875	,549		2,171	,044
Programacion	2,864	1,726	,394		1,660	,114
Java	-4,745	1,917	-,671		-2,475	,023
IDE	-12,385	10,850	-,207		-1,141	,269
TDD	13,857	5,352	,463		2,589	,019
5 (Constante)	35,591	10,432			3,412	,003
Academica	1,692	,863	,489		1,960	,065
Programacion	3,362	1,683	,462		1,997	,060
Java	-4,830	1,931	-,683		-2,502	,022
TDD	14,071	5,391	,470		2,610	,017

Tabla 5.22. Pasos de la stepwise regression para la calidad en el tratamiento TDD1

La Tabla 5.23 muestra el resumen de los modelos de regresión que en cada paso va arrojando la stepwise regression para la variable respuesta *Calidad*. En este caso son 5 modelos. El modelo elegido en este caso es el 5, la elección del modelo se ha realizado siguiendo el mismo criterio que en el apartado 5.4.1.1.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,699 ^a	,488	,215	20,19679
2	,699 ^b	,488	,264	19,56092
3	,697 ^c	,486	,305	19,01054
4	,681 ^d	,463	,314	18,87809
5	,652 ^e	,425	,303	19,02794

Tabla 5.23. Resumen de los modelos de regresión para la calidad en el tratamiento TDD1

En el modelo elegido las variables demográficas que son significativas son la *experiencia académica*, la *experiencia de programación*, la *experiencia en Java* y la *experiencia en TDD*.

5.4.2.3. Tratamiento TDD2

La Tabla 5.24 muestra los coeficientes (betas) de los modelos de regresión obtenidos para la variable respuesta *Calidad* correspondientes al tratamiento TDD2. Las variables demográficas que van mostrando significación a medida que avanza la stepwise regression son: la *Experiencia Académica*, la *Experiencia en Programación*, la *Experiencia en Java*, la *Experiencia en Testing* y la *Experiencia en TDD*.

La Tabla 5.25 muestra el resumen de los modelos de regresión que en cada paso va arrojando la stepwise regression para la variable respuesta *Calidad*. En este caso son 9 modelos. El modelo elegido en este caso es el 4, la elección del modelo se ha realizado siguiendo el mismo criterio que en el apartado 5.4.1.1.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,634 ^a	,402	,061	14,17849
2	,627 ^b	,393	,110	13,80011
3	,614 ^c	,377	,143	13,54282
4	,587 ^d	,345	,152	13,47011
5	,539 ^e	,291	,133	13,62439
6	,478 ^f	,229	,107	13,82505
7	,370 ^g	,137	,050	14,25915
8	,323 ^h	,104	,061	14,17569
9	,000 ⁱ	0,000	0,000	14,63162

Tabla 5.25. Resumen de los modelos de regresión para la calidad en el tratamiento TDD2

En el modelo elegido las variables significativas son la *experiencia en programación* y la *experiencia en TDD*.

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
1 (Constante)	52,704	12,415		4,245	,001
Academica	1,165	1,057	,516	1,102	,289
Programacion	-3,152	1,852	-,671	-1,702	,111
Java	1,604	2,050	,354	,782	,447
Testing	-3,387	2,167	-,621	-1,563	,140
JUnit	1,481	2,541	,250	,583	,569
IDE	4,095	8,934	,108	,458	,654
MetDes	6,359	10,088	,168	,630	,539
TDD	6,428	4,230	,336	1,520	,151
2 (Constante)	56,711	8,579		6,610	,000
Academica	1,301	,986	,577	1,319	,207
Programacion	-3,523	1,622	-,750	-2,172	,046
Java	1,758	1,968	,388	,893	,386
Testing	-3,596	2,062	-,659	-1,744	,102
JUnit	1,577	2,465	,266	,639	,532
MetDes	6,916	9,747	,183	,709	,489
TDD	6,443	4,117	,337	1,565	,138
3 (Constante)	55,325	8,146		6,791	,000
Academica	,896	,741	,397	1,208	,245
Programacion	-3,183	1,504	-,678	-2,116	,050
Java	2,375	1,684	,524	1,411	,177
Testing	-2,753	1,556	-,505	-1,769	,096
MetDes	8,402	9,290	,223	,904	,379
TDD	6,552	4,037	,343	1,623	,124
4 (Constante)	58,023	7,540		7,696	,000
Academica	1,139	,687	,505	1,658	,116
Programacion	-2,783	1,430	-,593	-1,946	,068
Java	1,886	1,586	,416	1,189	,251
Testing	-2,564	1,534	-,470	-1,672	,113
TDD	7,372	3,912	,385	1,884	,077
5 (Constante)	54,467	7,001		7,780	,000
Academica	1,310	,680	,580	1,927	,070
Programacion	-1,971	1,271	-,420	-1,551	,138
Testing	-1,719	1,374	-,315	-1,251	,227
TDD	8,101	3,908	,423	2,073	,053
6 (Constante)	54,526	7,104		7,675	,000
Academica	,932	,618	,413	1,509	,148
Programacion	-2,150	1,282	-,458	-1,678	,110
TDD	7,667	3,950	,401	1,941	,067
7 (Constante)	55,189	7,313		7,547	,000
Programacion	-,850	,978	-,181	-,869	,395
TDD	6,437	3,987	,336	1,615	,122
8 (Constante)	49,904	4,036		12,363	,000
TDD	6,170	3,952	,323	1,561	,133
9 (Constante)	54,196	3,051		17,764	,000

Tabla 5.24. Pasos de la stepwise regression para la calidad en el tratamiento TDD2

5.4.3. Influencia de las variables demográficas sobre TUS

En este apartado se realiza un análisis de la influencia de las variables demográficas sobre la variables respuesta *Productividad*, se muestra el modelo de regresión elegido en cada paso para cada uno de los tres tratamientos: TLD, TDD1 y TDD2.

5.4.3.1. Tratamiento TLD

La Tabla 5.26 muestra los coeficientes (betas) de los modelos de regresión obtenidos para la variable respuesta *TUS* para el tratamiento TLD. Las variables demográficas que van

mostrando significación a medida que avanza la stepwise regression son: la *Experiencia en JUnit*, la *Experiencia en Metodologías de Desarrollo* y la *Experiencia en TDD*.

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
1 (Constante)	4,806	2,227		2,158	,048
Academica	,093	,175	,214	,533	,602
Programacion	,039	,287	,042	,135	,895
Java	-,379	,396	-,426	-,959	,353
Testing	-,111	,405	-,104	-,274	,788
JUnit	,783	,469	,665	1,671	,115
IDE	,064	1,621	,008	,039	,969
MetDes	-3,670	1,937	-,486	-1,895	,078
TDD	1,213	,800	,321	1,516	,150
2 (Constante)	4,864	1,618		3,006	,008
Academica	,094	,168	,216	,561	,583
Programacion	,035	,264	,038	,133	,896
Java	-,377	,379	-,423	-,996	,334
Testing	-,113	,389	-,106	-,291	,774
JUnit	,782	,454	,664	1,725	,104
MetDes	-3,657	1,848	-,484	-1,979	,065
TDD	1,211	,773	,321	1,567	,137
3 (Constante)	4,970	1,368		3,633	,002
Academica	,098	,161	,225	,610	,550
Java	-,361	,348	-,405	-1,037	,314
Testing	-,127	,363	-,119	-,350	,730
JUnit	,795	,430	,675	1,849	,082
MetDes	-3,574	1,688	-,473	-2,117	,049
TDD	1,197	,743	,317	1,611	,126
4 (Constante)	4,922	1,328		3,707	,002
Academica	,070	,136	,161	,515	,613
Java	-,350	,338	-,393	-1,036	,314
JUnit	,705	,336	,599	2,097	,050
MetDes	-3,502	1,635	-,464	-2,142	,046
TDD	1,212	,724	,321	1,675	,111
5 (Constante)	4,949	1,301		3,805	,001
Java	-,221	,223	-,248	-,991	,334
JUnit	,633	,300	,538	2,111	,048
MetDes	-3,085	1,392	-,409	-2,215	,039
TDD	1,134	,694	,300	1,634	,119
6 (Constante)	4,710	1,278		3,687	,001
JUnit	,432	,221	,367	1,956	,065
MetDes	-3,206	1,387	-,425	-2,312	,032
TDD	1,147	,693	,304	1,654	,114
7 (Constante)	5,245	1,286		4,079	,001
JUnit	,509	,225	,432	2,265	,034
MetDes	-3,088	1,441	-,409	-2,143	,044

Tabla 5.26. Pasos de la stepwise regression para la TUS en el tratamiento TLD

La tabla 5.27 muestra el resumen de los modelos de regresión que en cada paso va arrojando la stepwise regression para la variable respuesta *TUS*. En este caso son 7 modelos. El modelo elegido en este caso es el 6, la elección del modelo se ha realizado siguiendo el mismo criterio que en el apartado 5.4.1.1.

En el modelo elegido las variables significativas son la *experiencia en JUnit* y la *experiencia en metodologías de desarrollo*.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,638 ^a	,407	,090	2,74093
2	,638 ^b	,407	,147	2,65404
3	,637 ^c	,406	,196	2,57622
4	,634 ^d	,402	,236	2,51265
5	,627 ^e	,393	,265	2,46360
6	,601 ^f	,362	,266	2,46251
7	,524 ^g	,274	,205	2,56234

Tabla 5.27. Resumen de los modelos de regresión para la TUS en el tratamiento TLD

5.4.3.2. Tratamiento TDD1

La Tabla 5.28 muestra los coeficientes (betas) de los modelos de regresión obtenidos para la variable respuesta *TUS* correspondientes al tratamiento TDD1.

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
1 (Constante)	4,986	3,195		1,561	,139
Academica	-,067	,251	-,107	-,265	,795
Programacion	,098	,411	,075	,238	,815
Java	,217	,568	,171	,383	,707
Testing	,438	,581	,286	,753	,463
JUnit	,085	,672	,050	,126	,901
IDE	-3,296	2,327	-,306	-1,417	,177
MetDes	2,156	2,779	,200	,776	,450
TDD	,449	1,148	,083	,391	,701
2 (Constante)	4,930	3,065		1,608	,127
Academica	-,085	,196	-,137	-,434	,670
Programacion	,108	,391	,082	,276	,786
Java	,254	,473	,199	,537	,599
Testing	,483	,446	,315	1,083	,295
IDE	-3,303	2,253	-,306	-1,466	,162
MetDes	2,227	2,637	,206	,844	,411
TDD	,465	1,106	,086	,421	,680
3 (Constante)	5,388	2,505		2,150	,046
Academica	-,080	,190	-,128	-,419	,681
Java	,323	,389	,254	,831	,417
Testing	,455	,422	,297	1,077	,296
IDE	-3,500	2,078	-,325	-1,684	,110
MetDes	2,536	2,322	,235	1,092	,290
TDD	,423	1,065	,079	,397	,696
4 (Constante)	5,601	2,389		2,344	,031
Academica	-,098	,180	-,157	-,544	,593
Java	,361	,369	,284	,980	,340
Testing	,453	,412	,296	1,098	,286
IDE	-3,525	2,028	-,327	-1,738	,099
MetDes	2,697	2,232	,250	1,208	,243
5 (Constante)	5,732	2,333		2,458	,024
Java	,268	,321	,211	,837	,413
Testing	,397	,392	,259	1,014	,324
IDE	-3,709	1,962	-,344	-1,891	,074
MetDes	2,250	2,037	,209	1,105	,283
6 (Constante)	5,821	2,313		2,517	,020
Testing	,620	,285	,405	2,172	,042
IDE	-3,587	1,942	-,333	-1,847	,080
MetDes	2,300	2,020	,213	1,139	,268
7 (Constante)	7,296	1,929		3,782	,001
Testing	,711	,276	,465	2,577	,018
IDE	-3,343	1,944	-,310	-1,720	,100
8 (Constante)	4,483	1,068		4,196	,000
Testing	,721	,288	,471	2,505	,020

Tabla 5.28. Pasos de la stepwise regression para la TUS en el tratamiento TDD1

Las variables demográficas que van mostrando significación a medida que avanza la stepwise regression son: la *Experiencia Académica*, la *Experiencia en Testing*, la *Experiencia en IDE* y la *Experiencia en Metodologías de Desarrollo*.

La Tabla 5.29 muestra el resumen de los modelos de regresión que en cada paso va arrojando la stepwise regression para la variable respuesta *TUS*. En este caso son 8 modelos. El modelo elegido en este caso es el 6, la elección del modelo se ha realizado siguiendo el mismo criterio que en el apartado 5.4.1.1.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,634 ^a	,401	,082	3,93301
2	,633 ^b	,401	,139	3,81014
3	,631 ^c	,398	,185	3,70517
4	,626 ^d	,392	,224	3,61747
5	,618 ^e	,382	,252	3,54982
6	,600 ^f	,360	,263	3,52313
7	,564 ^g	,318	,253	3,54791
8	,471 ^h	,222	,187	3,70249

Tabla 5.29. Resumen de los modelos de regresión para la *TUS* en el tratamiento TDD1

En el modelo elegido las variables significativas son la *experiencia en Testing* y la *experiencia en IDE*.

5.4.3.3. Tratamiento TDD2

La Tabla 5.30 muestra los coeficientes (betas) de los modelos de regresión obtenidos para la variable respuesta *TUS* correspondientes al tratamiento TDD2. La única variables que va mostrando significación a medida que avanza la stepwise regression es la *Experiencia en Metodologías de Desarrollo*.

La Tabla 5.31 muestra el resumen de los modelos de regresión que en cada paso va arrojando la stepwise regression para la variable respuesta *TUS*. En este caso son 9 modelos. En este caso no se elige ningún modelo porque ninguna variable demográfica resulta significativa en ninguno de los modelos que se han generado en cada paso de la regresión. Por lo tanto se puede concluir que ninguna de las variables demográficas influye en la variable respuesta *TUS* para el tratamiento TDD2.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,312 ^a	,097	-,419	2,10886
2	,311 ^b	,097	-,325	2,03795
3	,305 ^c	,093	-,247	1,97737
4	,296 ^d	,087	-,181	1,92396
5	,288 ^e	,083	-,121	1,87408
6	,285 ^f	,081	-,064	1,82581
7	,284 ^g	,080	-,011	1,78055
8	,253 ^h	,064	,020	1,75291
9	,000 ⁱ	0,000	0,000	1,77042

Tabla 5.31. Resumen de los modelos de regresión para la *TUS* en el tratamiento TDD2

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
1 (Constante)	6,657	1,847		3,605	,003
Academica	,054	,157	,198	,344	,736
Programacion	-,080	,276	-,141	-,292	,775
Java	-,028	,305	-,050	-,090	,929
Testing	-,185	,322	-,281	-,575	,574
JUnit	,137	,378	,192	,364	,722
IDE	-,345	1,329	-,076	-,260	,799
MetDes	-1,014	1,500	-,222	-,676	,510
TDD	-,141	,629	-,061	-,224	,826
2 (Constante)	6,685	1,760		3,799	,002
Academica	,048	,136	,175	,351	,731
Programacion	-,088	,253	-,155	-,349	,732
Testing	-,187	,311	-,283	-,600	,558
JUnit	,121	,323	,169	,376	,712
IDE	-,365	1,267	-,080	-,288	,777
MetDes	-,959	1,329	-,210	-,722	,481
TDD	-,150	,600	-,065	-,251	,805
3 (Constante)	6,623	1,691		3,918	,001
Academica	,053	,131	,193	,404	,692
Programacion	-,092	,245	-,162	-,375	,713
Testing	-,184	,302	-,279	-,609	,551
JUnit	,111	,310	,154	,356	,726
IDE	-,376	1,228	-,082	-,306	,763
MetDes	-1,011	1,274	-,221	-,794	,439
4 (Constante)	6,256	1,161		5,388	,000
Academica	,037	,117	,135	,315	,756
Programacion	-,061	,217	-,107	-,281	,782
Testing	-,165	,287	-,249	-,573	,574
JUnit	,093	,296	,129	,312	,759
MetDes	-1,037	1,236	-,227	-,838	,413
5 (Constante)	6,080	,953		6,382	,000
Academica	,016	,089	,060	,184	,856
Testing	-,137	,263	-,207	-,521	,609
JUnit	,046	,239	,064	,191	,851
MetDes	-1,073	1,198	-,235	-,896	,382
6 (Constante)	6,092	,926		6,579	,000
Testing	-,108	,206	-,164	-,525	,606
JUnit	,032	,221	,044	,144	,887
MetDes	-,979	1,055	-,214	-,928	,365
7 (Constante)	6,088	,903		6,745	,000
Testing	-,088	,148	-,134	-,594	,559
MetDes	-,972	1,028	-,213	-,946	,356
8 (Constante)	6,000	,876		6,846	,000
MetDes	-1,158	,964	-,253	-1,201	,243
9 (Constante)	5,043	,369		13,662	,000

Tabla 5.30. Pasos de la stepwise regression para la TUS en el tratamiento TDD2

5.4.4. Resumen de resultados

La Tabla 5.32 muestra un resumen de los resultados que se han presentado en los apartados 5.4.1, 5.4.2 y 5.4.3. Por cada variable respuesta (ej. Productividad, Calidad, TUS) se presenta el modelo que se ha elegido en la stepwise regression. Una x indica que la variable no pertenece al modelo de regresión. Cuando una variable pertenece al modelo se indica su p-valor. Los p-valores se han coloreado en rojo o verde para denotar significación estadística al 5% y 10%, respectivamente.

Las variables demográficas experiencia en JUnit y experiencia en metodologías de desarrollo son las más influyente en la productividad, dado que son significativas en más de

un tratamiento, en menor medida influyen también la experiencia en *Testing* y la experiencia en *IDE*. Con el mismo criterio se puede afirmar que la experiencia de *programación* y la experiencia en *TDD* son las más influyentes en la *Calidad*, y en menor medida la experiencia *académica*, la experiencia en *Java* y la experiencia en *JUnit*. En el caso la variable respuesta *TUS*, tiene ligera influencia de las variables experiencia en *Testing*, en *JUnit*, *IDE* y *metodologías de desarrollo*.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	x	x	0,177	x	0,065	0,116	x	x	x
Programación	x	x	0,282	x	0,06	0,068	x	x	x
Java	x	x	x	x	0,022	0,251	x	x	x
Testing	x	0,157	0,017	x	x	0,113	x	0,042	x
JUnit	0,069	x	0,046	0,088	x	x	0,065	x	x
IDE	x	0,098	x	x	x	x	x	0,08	x
MetDes	0,013	x	0,089	0,243	x	x	0,032	0,268	0,243
TDD	0,152	x	x	0,081	0,017	0,077	0,114	x	x

Tabla 5.32. p-valores del modelo de regresión elegido por cada variable respuesta y tratamiento

Las betas para las variables demográficas que forman parte del modelo de regresión en cada una de las variables respuesta son las que se muestran en la Tabla 5.33. Se usa el mismo criterio de colores que la Tabla 5.32, la única diferencia es que aquí se muestran las betas del modelo de regresión en lugar del p-valor. El coeficiente de determinación R^2 se muestra en negrita en la primera fila.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
R^2	,379	,200	,423	,304	,425	,345	,362	,360	,064
Académica	x	x	,53	x	1,69	1,13	x	x	x
Programación	x	x	-,77	x	3,36	-2,78	x	x	x
Java	x	x	x	x	-4,83	1,88	x	x	x
Testing	x	3,34	-2,46	x	x	-2,56	x	,620	x
JUnit	2,74	x	2,06	4,83	x	x	,432	x	x
IDE	x	-27,73	x	x	x	x	x	-3,58	x
MetDes	-24,12	x	-7,19	-20,35	x	x	-3,20	2,30	-1,1
TDD	6,61	x	x	15,55	14,07	7,37	1,14	x	x

Tabla 5.33. Coeficientes (betas) del modelo de regresión elegido por cada variable respuesta y tratamiento

Otra vez sorprenden los resultados que se obtienen. Por ejemplo la experiencia en *Java* que tienen los sujetos, influye negativamente en la *Calidad*. Lo mismo sucede con la experiencia en *metodologías de desarrollo*. Lo que lleva a pensar que la muestra con los que se cuenta para hacer el análisis no es lo suficientemente amplia como para realizar un análisis demográfico.

CAPÍTULO 6. DISCUSIÓN

Este capítulo se presenta la discusión acerca de los resultados obtenidos en el capítulo 5: Resultados. Primeramente se realiza una comparación entre los resultados obtenidos en el análisis dicotómico y los resultados del análisis de correlación de Pearson. A continuación se comparan los resultados que se obtienen en la dicotomización y en el análisis de correlación de Pearson frente a los resultados que se obtienen en el análisis de Regresión Lineal. Finalmente se hace una síntesis general de los resultados obtenidos en todos los análisis y se comenta la fiabilidad que creemos tienen los resultados.

6.1. Dicotomización vs Correlación de Pearson

Para realizar este análisis, se compara la significación del test de comparación de medias que se obtuvo en el análisis dicotómico (T-Student y U de Mann-Whitney), con la significación obtenida por el test de correlación de Pearson. En primera instancia se presentan tablas resumen de la comparación de los datos dicotomizados y los obtenidos en la correlación. A continuación se analizarán en detalle las coincidencias y discrepancias. Este estudio se estructura en tres apartados:

- El primer lugar se estudia el caso en el que los dos test (comparación de medias y correlación) dan resultado significativo, es decir, el resultado de los dos test coinciden.
- Segundo, se comparan los casos en los que en el test de comparación de medias se obtiene un valor significativo y no así en la correlación de Pearson.
- Tercero, se comparan los casos en los que saliendo significativo en la correlación, no se obtiene el mismo resultado en la comparación de medias con los datos dicotomizados.

En cada apartado se analiza el diagrama de cajas que se ha obtenido con los datos dicotomizados, frente al gráfico de dispersión que se obtiene con los datos no dicotomizados, para comprobar si son coherentes. Finalmente se presenta unas conclusiones preliminares sobre la fiabilidad de los resultados que se obtienen tras realizar estos dos tipos de análisis (dicotomización y correlación de Pearson).

6.1.1. Resumen de similitudes y diferencias

La Tabla 6.1 muestra las discrepancias y similitudes de los resultados de los tests de comparación de medias y correlación de Pearson. Un ✓ indica que el resultado de los dos test coincide, mientras que una X indica que hay discrepancia entre el resultado de los mismos.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	✓	X	✓	✓	X	✓	X	X	✓
Programación	✓	X	✓	✓	✓	✓	✓	✓	✓
Java	✓	✓	✓	✓	✓	✓	✓	X	✓
Testing	X	X	X	X	X	✓	X	X	✓
JUnit	✓	✓	✓	X	✓	✓	X	X	✓
IDE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MetDes	X	✓	X	✓	✓	✓	✓	✓	✓
TDD	✓	✓	✓	X	✓	✓	X	✓	✓

Tabla 6.1. Similitudes y diferencias entre los resultados de la dicotomización y la correlación

La discrepancia de los resultados de la comparación de medias con los datos dicotomizados y la correlación de Pearson es bastante alta, mucho más de lo que podría esperarse por puro error experimental. En los siguientes apartados se discutirán las similitudes y discrepancias.

6.1.2. Dicotomización significativa vs Correlación NO significativa

A continuación estudiaremos los casos en los que no coinciden los resultados que arrojan los dos tipos de análisis (comparación de medias y correlación de Pearson). A modo de ejemplo se toman tres pares de variables del tratamiento TLD. La *experiencia en Testing* frente a las tres variables respuesta: *Productividad*, *Calidad* y *TUS*. En la figura 6 se muestra el diagrama de de cajas y el gráfico de dispersión para los tres pares de variables.

Analizando los diagramas de cajas y los gráficos de dispersión, se puede notar que, las diferencias en los resultados de los test pueden deberse a un efecto pernicioso de la dicotomización. Esto se debe a que la división en grupos durante el proceso de dicotomización se realiza de manera arbitraria.

Otro motivo por el que se presentan estas discrepancias es la presencia de outliers. Los outliers hacen que el poder de la correlación disminuya considerablemente provocando

que los valores altos influyan considerablemente la pendiente de la recta que presenta el ajuste lineal.

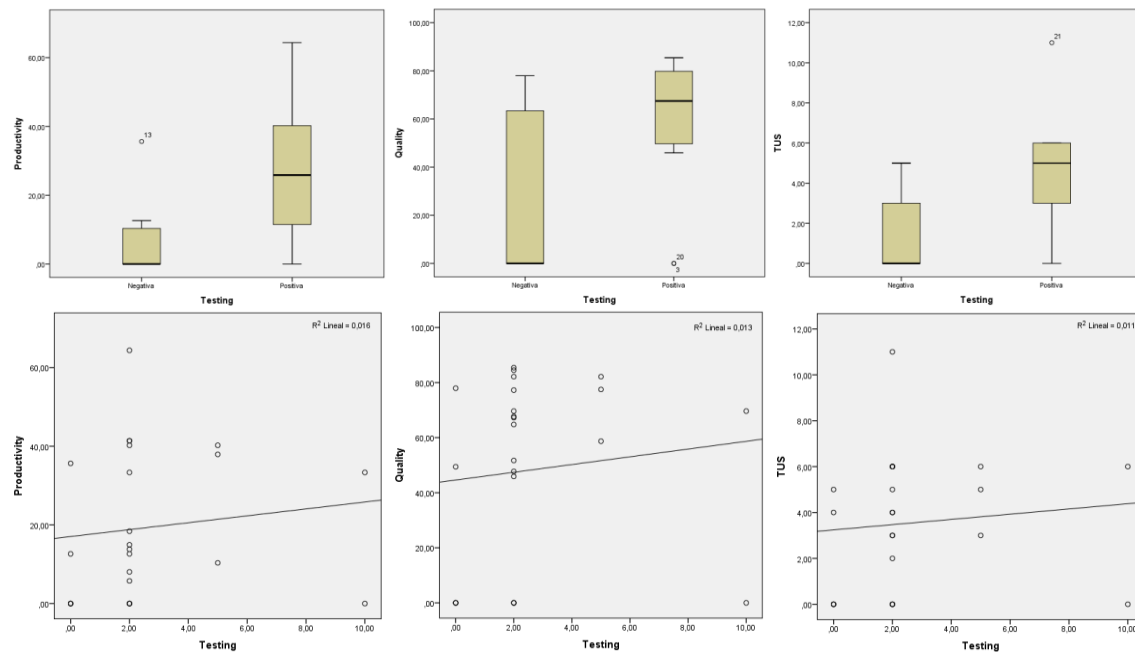


Figura 6. Diagramas de cajas y gráficos de dispersión: experiencia en Testing vs (Productividad, Calidad, TUS)

El análisis dicotómico muestra una clara influencia de la experiencia en *Testing* de los sujetos en la *Productividad*, sin embargo no se obtiene el mismo resultado del análisis de correlación de Pearson. Con los datos cuantitativos se obtiene una representación más fiel y permite ver con más detalle que es lo que realmente está sucediendo.

6.1.3. Correlación significativa vs Dicotomización NO significativa

En este apartado se estudian los casos en los que no coinciden los resultados que arrojan los dos tipos de análisis (comparación de medias y correlación de Pearson), en este caso, el test de correlación indica que las variables están relacionadas mientras que el test de comparación de medias no encuentra evidencias para afirmar que exista relación. Se toma tres pares de variables de tratamiento TDD1. La *experiencia en Java* frente a *TUS*, La *experiencia en Testing* frente a *TUS* y la *experiencia en JUnit* frente a *TUS*. En la Figura 7 se muestra el diagrama de de cajas y el grafico de dispersión para los tres pares de variables.

El motivo por el que el test de comparación de medias con los datos dicotomizados afirma que no hay diferencia considerable en las medias de las dos categorías (Negativa, Positiva), es el problema de la arbitrariedad en la definición de subgrupos durante el proceso de dicotomización de los datos. Los diagramas de cajas no parecen mostrar diferencias entre grupos, esto se debe a que la caja engloba puntos de datos con mucha diversidad.

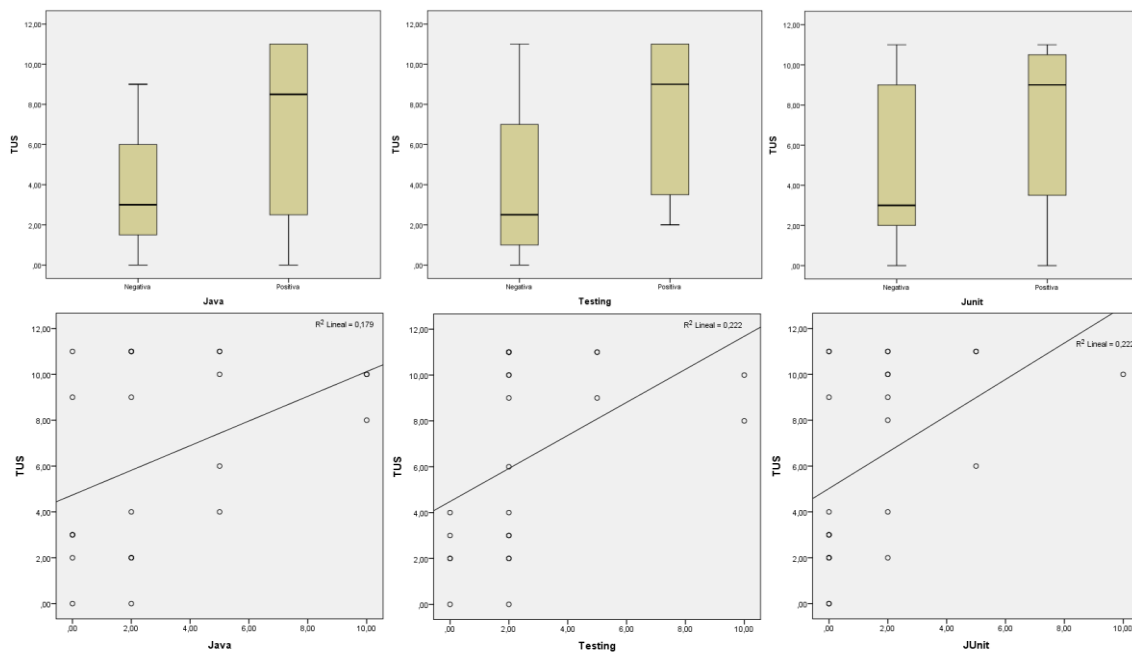


Figura 7. Diagramas de cajas y gráficos de dispersión: (experiencia en Java, Testing, JUnit) vs TUS

6.1.4. Sobre los resultados

El análisis que se ha llevado a cabo en los apartados anteriores se realiza para todos los pares de variables que resultan de combinar, las variables demográficas, con las tres variables respuesta del experimento (Productividad, Calidad, TUS), por cada una de los tratamientos: TLD, TDD1 y TDD2.

Tras el análisis comparativo de los resultados obtenidos en el análisis dicotómico frente a los datos obtenidos en el análisis de correlación de Pearson, se puede afirmar que con la dicotomización se pierde mucha información que puede ser de suma importancia para los resultados del análisis. El análisis completo que apoya estos resultados se encuentra en el Anexo D. Comparación de la dicotomización y la correlación de Pearson.

Asimismo, el pequeño tamaño de la muestra (reducido número de sujetos experimentales) resulta un problema, ya que tras la división en grupos en el proceso de dicotomización, en algunos casos uno de los grupos tiene muy pocos sujetos, esto hace que los resultados de los test sean en algunos casos producto de la casualidad.

También es un problema la presencia de outliers. Los diagramas de cajas engloban datos con mucha diversidad provocando que no muestren diferencias entre los grupos cuando realmente las hay. Por otro lado, los outliers afectan directamente a los resultados de la correlación de Pearson.

En general los diagramas de cajas muestran diferencia entre los grupos, sin embargo al analizar los gráficos de dispersión se puede vislumbrar que no existen tales diferencias. Los

datos cuantitativos aportan mayor información, lo que permite aclarar situaciones que con la dicotomización no es posible.

6.2. Regresión lineal vs Dicotomización

Con el fin de afinar más los resultados obtenidos con el análisis dicotómico y el análisis de correlación de Pearson, se realizó un análisis de regresión lineal múltiple. A continuación se comparan los resultados del análisis de regresión lineal y la dicotomización. La Tabla 6.2 muestra las similitudes y discrepancias de los dos tipos de análisis. La interpretación de la tabla es la misma que la de la Tabla 6.1.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	✓	X	✓	✓	✓	✓	X	X	✓
Programación	✓	X	✓	✓	✓	✓	✓	X	✓
Java	✓	✓	✓	✓	X	✓	✓	✓	✓
Testing	X	X	X	X	X	✓	X	✓	✓
JUnit	✓	✓	✓	✓	✓	✓	✓	✓	✓
IDE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MetDes	X	✓	✓	✓	✓	✓	X	✓	✓
TDD	✓	✓	✓	✓	✓	✓	✓	✓	✓

Tabla 6.2. Similitudes y diferencias entre los resultados de la regresión lineal y la dicotomización

Se encuentran 14 discrepancias entre los dos análisis. Dada las discrepancias que se obtuvieron al comparar los resultados de la dicotomización y la correlación de Pearson, este resultado no sorprende. La regresión lineal es un análisis que generaliza un análisis de correlación, con lo cual los resultados de la regresión lineal deben parecerse más a los resultados de la correlación. Esto se analiza en el siguiente apartado.

6.3. Regresión lineal vs Correlación de Pearson

A continuación se comparan los resultados del análisis de regresión lineal y la correlación de Pearson. La Tabla 6.3 muestra las similitudes y discrepancias de los dos tipos de análisis. La interpretación de la tabla es la misma que la de la tabla 6.1.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	✓	✓	✓	✓	X	✓	✓	✓	✓
Programación	✓	✓	✓	✓	✓	✓	✓	X	✓
Java	✓	✓	✓	✓	X	✓	✓	X	✓
Testing	✓	✓	✓	✓	✓	✓	✓	X	✓
JUnit	✓	✓	✓	X	✓	✓	✓	X	✓
IDE	✓	X	✓	✓	✓	✓	✓	✓	✓
MetDes	✓	✓	X	✓	✓	✓	X	✓	✓
TDD	✓	✓	✓	X	✓	✓	X	✓	✓

Tabla 6.3. Similitudes y diferencias entre los resultados de la regresión lineal y la correlación

Se esperaba que no hubiera mucha discrepancia entre los dos tipos de análisis puesto que la regresión lineal está muy relacionada con la correlación de Pearson.

Hay 12 discrepancias, más de lo que cabría esperar dada la similitud entre la correlación y la regresión lineal múltiple, aunque tampoco resulta sorprendente.

Los resultados que se obtienen en la regresión lineal son más fiables que los que se obtiene en la correlación de Pearson, esto se debe a que cuando se hace una regresión se tienen en cuenta todas las variables a la hora de realizar el modelo, mientras que en el caso de la correlación solo se compara una variable con otra sin tener en cuenta el resto de variables. Las variables demográficas podrían estar relacionadas de alguna forma unas con otras, por ejemplo es posible que una persona que tiene muchos años de experiencia de programación presente mayor conocimiento sobre metodologías de desarrollo. Esto explica la diferencia en los resultados.

El análisis de correlación lineal es un método menos sofisticado que la regresión lineal múltiple, puesto que en la regresión lineal múltiple se tiene en cuenta todas las variables, por lo tanto resulta ser un método más completo y confiable que una correlación bivariada.

6.4. Regresión lineal: entrada forzosa vs stepwise regression

En este apartado se compara la regresión con entrada forzosa frente a la regresión paso a paso hacia atrás (stepwise regression). La Tabla 6.4 muestra las similitudes y discrepancias de los dos tipos de análisis. Un guion (-) indica que esa variable demográfica no entro en el modelo elegido de la stepwise regression. Un signo ✓ independientemente del color, indica que esa variable entro en el modelo elegido de la stepwise regression. Cuando el signo ✓ es de color rojo (✓), significa que esa variable no resulto significativa en la regresión lineal multivariante pero sí en la stepwise regression.

Variable demográfica	Productividad			Calidad			TUS		
	TLD	TDD1	TDD2	TLD	TDD1	TDD2	TLD	TDD1	TDD2
Académica	-	-	✓	-	✓	✓	-	-	-
Programación	-	-	✓	-	✓	✓	-	-	-
Java	-	-	-	-	✓	✓	-	-	-
Testing	-	✓	✓	-	-	✓	-	✓	-
Junit	✓	-	✓	✓	-	-	✓	-	-
IDE	-	✓	-	-	-	-	-	✓	-
MetDes	✓	-	✓	✓	-	-	✓	✓	-
TDD	✓	-	-	✓	✓	✓	✓	✓	-

Tabla 6.4. Similitudes y diferencias entre los resultados de la regresión lineal y la stepwise regression

Todas las variables demográficas que son significativas en la regresión lineal múltiple también lo son en la stepwise regression. Con lo cual, el análisis de regresión múltiple está totalmente incluido en la stepwise regression ya que todas las variables que resultaron significativas en la regresión lineal múltiple también los son en la stepwise regression.

Como se puede apreciar la stepwise regression resulta ser más completa que la regresión lineal múltiple, y es capaz de predecir con mayor exactitud.

6.5. Confiabilidad de los resultados obtenidos en los análisis

La dicotomización no es el mejor camino a seguir cuando se realizan un análisis Post-Hoc [3][4]. El principal problema de la dicotomización es que la clasificación en categorías se realiza de manera arbitraria, sin seguir ningún criterio matemático. Por lo tanto, resulta una clasificación generalmente subjetiva, es decir, depende del criterio del investigador. Esto es así salvo casos en los que resulta muy evidente la división en dos grupos (ej. hombre y mujer). La dicotomización presenta además un problema de pérdida de información. Con la dicotomización de variables cuantitativas se pierde parte de la información que puede provocar un sesgo en las estimaciones de la significación estadística y del tamaño de efecto. Estos problemas se solucionan en parte realizando el análisis con los datos cuantitativos utilizando, por ejemplo, la correlación de Pearson.

Con respecto a la correlación se puede decir que aunque mejora el análisis con datos dicotomizados sigue sin ser la mejor solución. Presenta varios problemas. Uno de ellos es que las comparaciones se realizan por pares, por lo tanto no se tiene en cuenta la posible relación que puede haber entre las variables demográficas. Este problema lo resuelve la regresión lineal. Otro problema que presenta el poder de la correlación se ve muy afectada por la presencia de outliers.

El tamaño de la muestra resulta ser un problema que afecta a los cuatro tipos de análisis que se han realizado. En el caso de la dicotomización, se ve afectada a la hora de dividir en grupos. Cuando uno de los grupos es muy pequeño los resultados no son fiables. La regresión lineal también se ve afectada por el tamaño de la muestra. Para detectar una correlación con un bajo tamaño de efecto (según Cohen) serían necesarias alrededor de 600 muestras [7].

Los resultados que se obtienen con la dicotomización y la correlación de Pearson se mejoran con el análisis de regresión lineal. Sin embargo el coeficiente de determinación o bondad de ajuste que se obtiene con la regresión lineal tampoco es muy buena, debido probablemente en parte al tamaño muestral [8].

La ventaja que ofrece la stepwise regression es que resuelve los posibles problemas de multicolinealidad que puede existir entre las variables. Esto es gracias a que no mantiene fijas las variables que pertenecen al modelo en cada etapa de la stepwise regression

CAPÍTULO 7. CONCLUSIONES

Este TFM tiene dos objetivos principales. El primer objetivo consiste en analizar la influencia de las variables demográficas en las variables respuesta de un experimento acerca de TDD. Después de analizar los resultados obtenidos con diferentes técnicas de análisis, se puede concluir que en términos generales no se han encontrado evidencias suficientes como para afirmar que las variables demográficas influyan sustancialmente en los resultados del experimento. Sin embargo, sí se encuentran resultados puntuales en los que parece que hay indicios de dicha influencia, pero no se han considerado suficientes como para realizar dicha afirmación. Por ejemplo, la experiencia en JUnit parece tener influencia positiva sobre la Productividad y la experiencia en TDD sobre la calidad. No obstante, a estos resultados no se le puede dar mucha credibilidad dado que la bondad de ajuste los tests que muestran estas evidencias es bastante baja.

Algunos de los resultados obtenidos no están muy acorde con lo que indica el sentido común. Por ejemplo se tiende a pensar que alguien con mucha experiencia debería reportar una mayor *Calidad* y *Productividad*. Sin embargo los datos no muestran eso precisamente. Por ejemplo en el caso del sujeto FSK06 no presenta *Experiencia en TDD*, pero obtiene una calidad y productividad por encima de la media. Esto no es tan descabellado si se piensa que el experimento se ejecuto en el contexto empresarial, y que por ejemplo a un sujeto con 30 años de experiencia de programación, que usa una técnica de programación que no es con la que se realiza el experimento, le pueda costar más aprender una técnica nueva. Se puede dar también por ejemplo que alguien con poca experiencia de programación obtenga mejor productividad o calidad, dado que asimila mejor nuevos conceptos.

Finalmente, los resultados que se obtienen en los análisis sugieren que para poder realizar un estudio de la información demográfica de los sujetos experimentales, se requiere que la muestra sea más grande.

El segundo objetivo consiste en analizar la influencia de la metodología usada para el análisis de los datos. Se realizaron cuatro tipos de análisis: comparación de medias con

datos dicotomizados, correlación de Pearson, regresión lineal multivariante y stepwise regression.

Se puede decir que aunque el ser humano este acostumbrado a tratar con datos dicotomizados, este no es el camino para realizar un análisis Post-Hoc, dado que se pierde información importante en el proceso de dicotomización de los datos [3].

La correlación de Pearson solventa el problema de la división en grupos de la dicotomización. No obstante también presenta algunos problemas, uno de ellos es que las comparaciones se realizan por pares, dado que no se tiene en cuenta la posible relación que puede existir entre las variables demográficas. Otro problema que presenta el poder de la correlación se ve muy afectada por la presencia de outliers.

La regresión lineal evita los problemas de la comparación por pares de la correlación, ya que tiene en cuenta todas las variables a la vez. Sin embargo tenemos el problema de que la muestra es pequeña como para que el resultado del análisis sea fiable. El tamaño de la muestra resulta determinante a la hora de realizar este análisis. Influye directamente sobre la bondad de ajuste de los modelos que se obtienen.

La stepwise regression tiene la ventaja de no mantener fijas las variables que ya entraron en el modelo en un paso de la regresión, evitando de esta forma los problemas de multicolinealidad [8].

Queda como principal trabajo futuro realizar este análisis con un tamaño muestral más representativo, para poder responder con más certeza acerca de la influencia de las variables demográficas sobre los experimentos de Ingeniería del Software.

REFERENCIAS

- [1] Juristo N., Moreno A., Basics of Software Engineering Experimentation, Spain. 2001
- [2] Tore Dybå, Dag I. K. Sjøberg, and Daniela S. Cruzes: What works for whom, where, when, and why? On the Role of Context in Empirical Software Engineering. ESEM 2012: 19-28
- [3] MacCallum, R. C., S. Zhang, K. J. Preacher, and D. D. Rucker. 2002. On the Practice of Dichotomization of Quantitative Variables. *Psychological Methods* 7 (1): 19-40.
- [4] Marin F., Sánchez J., and Huedo T. The cost of Dichotomizing in Meta-Analysis. *Dpto Psicología Básica y Metodología, Universidad de Murcia*. 2003.
- [5] Razali N., Wah Y 2011 Power Comparison of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson Darling tests. *Journal of Statistical Modeling and Analytics*
- [6] McDonald, J.H. 2009. Handbook of Biological Statistics (2nd ed.). *Sparky House Publishing*, Baltimore, Maryland. 256-259.
- [7] Andy Field; Jeremy Miles; Zoe Field. Discovering Statistics Using R. SAGE. 2012.
- [8] E. López Gonzales. Tratamiento de la Colinealidad en Regresión Múltiple. *Psicothema*, 1998 vol 10, nº 2 pp.491 – 507.

Recursos Online:

- [1] Estadística para investigadores. Available: <http://diarium.usal.es/mooc/estadistica-para-investigadores/>
- [2] Rules of thumb on magnitudes of effect sizes. Available: <http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize>

ANEXO A. Análisis Demográfico

1 EJECUCIÓN FSECURE OULU

Este apartado contiene un análisis de los resultados de los formularios demográficos que reportaron los sujetos de la empresa Oulu.

1.1 FORMACIÓN ACADÉMICA

Las calificaciones de los sujetos se muestran en la Tabla 1. Todos ellos tienen títulos universitarios. Seis de los siete sujetos tienen formación universitaria a nivel de máster. El sujeto S03 intentó estudiar el máster, pero no se graduó. Ninguno de los sujetos, con la excepción del tema S05 (Certificado por Sun), logró una certificación profesional.

La experiencia en general de los sujetos también se muestra en la Tabla 1. Tienen más de 10 años de experiencia en promedio, oscila entre 7,5 y 30 años. Todos menos S06 poseen una amplia experiencia en desarrollo, que oscila entre 6 y 30 años. Sus cargos actuales parecen estar relacionados con el desarrollo. S06 no es un desarrollador, sin embargo tiene casi 9 años de experiencia en Testing.

Tabla 1. Cualificación y experiencia en general de los sujetos

ID	Academic Degree	Current position	Professional career	Professional certifications
S01	MSc in Mathematics	Software Engineer	Developer: 14 years	
S02	MSc in Computer Engineering	Senior Software Engineer	Researcher: 5 years Developer: 6 years	
S03	MSc in Computer Science (not graduated)	R&D, Software engineer	Developer: 30 years, being also project manager Product manager: short times at a time	
S04	MSc in Computer Science	Senior Software Engineer	Developer: 10 years	
S05	MSc in Computer Science	Senior Software Engineer	Developer: 7.5 years	Sun certified Java Programmer Sun certified Java Business component developer
S06	MSc in Electrical Engineering	Quality Engineer	Tester: 6 years Quality engineer: 2,5 years	
S07	MSc in Electrical Engineering	Developer	Developer: 10 years	

1.2 EXPERIENCIA DE PROGRAMACIÓN

Los sujetos han utilizado varios lenguajes de programación, durante su educación y, sobre todo, durante sus carreras profesionales. Como se muestra en la Tabla 2, los lenguajes que se utilizan con más frecuencia son Java, Python y C (en diferentes versiones). S06 representa de nuevo una excepción, porque reporta únicamente experiencia en Python.

Tabla 2. Experiencia en lenguaje de programación

ID	Most frequently used			Intermediate			Least frequently used		
	P.L.	Educ. /yr	Usage /yr.	P.L.	Educ. /yr	Usage /yr.	P.L.	Educ. /yr	Usage /yr.
S01	Java	1	8	PL/SQL	1	5	Python	0	1
S02	Python	0	4	C++	0	5	Java	0	2
S03	Java	0	14	C	0	7	C++	0	2.5
S04	Java	3	5	Python	0	1	C#	0	1
S05	Java	8	7	Python	0	2	C	1	0
S06	Python	0	3	-	-	-	-	-	-
S07	C++	0	6	Python	0	2	Java	1	0

La falta de programación y / o experiencia en Java, como el caso de S06, representa una amenaza para el experimento debido a tareas de codificación se hará y se utilizarán lenguaje Java. Para corroborar que los sujetos eran competentes en la programación y Java, que independientemente les preguntamos acerca de los dos aspectos. Los resultados se muestran en la Tabla 3. Los datos reportados en la Tabla 2 y en la Tabla 3 son consistentes en general, aunque hay excepciones. Por ejemplo, S04 sobreestima su experiencia en programación, es decir: se informa a sí mismo como "Experto (> 10 años)" cuando, de acuerdo con la Tabla 2, que tiene sólo 7 años de experiencia. S06 reporta poca experiencia en programación, ninguno con el lenguaje de programación Java.

Tabla 3. Experiencia de programación y Java

ID	Programming experience	Java experience
S01	Expert (>10 years)	Intermediate (5-<=10 years)
S02	Expert (>10 years)	Novice (2-<=5 years)
S03	Expert (>10 years)	Expert (>10 years)
S04	Expert (>10 years)	Intermediate (5-<=10 years)
S05	Intermediate (5-<=10 years)	Intermediate (5-<=10 years)
S06	Novice (2-<=5 years)	No experience (<2 years)
S07	Intermediate (5-<=10 years)	Novice (2-<=5 years)

1.3 EXPERIENCIA EN TESTING

La experiencia en testing puede favorecer la aplicación de TDD. Por esta razón, conocer la experiencia de las pruebas unitarias de los sujetos puede ser útil para explicar los valores atípicos y / o comparar los resultados de la ejecución actual con otras ejecuciones. La Tabla 4 presenta los resultados. Todas las asignaturas tienen una experiencia limitada (de principiante-intermedio), con excepción de S03. S03 reporta más de 10 años de experiencia en pruebas de unidad, lo que coincide con sus 30 años de experiencia en el desarrollo.

Tabla 4. Experiencia en pruebas de unidad

ID	Unit testing experience
S01	Novice (2-<=5 years)
S02	Intermediate (5-<=10 years)
S03	Expert (>10 years)
S04	Intermediate (5-<=10 years)
S05	Intermediate (5-<=10 years)
S06	Novice (2-<=5 years)
S07	Novice (2-<=5 years)

1.4 EXPERIENCIA EN FRAMEWORKS DE TESTING

La aplicación de TDD requiere la disponibilidad de infraestructuras de prueba para realizar el ciclo de test-fail-code-success-refactor. Por lo tanto, la experiencia previa en frameworks de testing, en particular JUnit (que se utilizará en la formación y tareas experimentales), puede ejercer una influencia positiva en los resultados experimentales. Alternativamente, la falta de experiencia puede causar efectos adversos.

Como se muestra en la Tabla 5, todos los sujetos informan de experiencia en el uso de algún framework de testing, en algunas ocasiones (S01, S02, S04 y S05) más de uno. Los sujetos S01-S05 han utilizado JUnit durante al menos 2 años. S04 y S05 tienen aún más experiencia, reportando un uso de más de 5 años.

S06 y S07 (resaltado en el cuadro 5) no han utilizado JUnit, sin embargo han usado el módulo Python unittest. Esto no representa un inconveniente, al menos en principio, porque tal módulo, también denominado PyUnit, es la versión de Python de JUnit. Por lo tanto, la experiencia unittest podría transferirse fácilmente a JUnit. El lado negativo es que S06 y S07 tienen una experiencia limitada (<2 años) en la utilización de unittest.

Parece razonable suponer que la experiencia con los marcos de pruebas puede ser una variable moderadora, con 2 niveles: sin experiencia y con experiencia. Otra gradación (por ejemplo: novato vs intermedio) podría ayudar a afinar el análisis.

Tabla 5. Experiencia en frameworks de testing y JUnit

ID	Unit test tools used	JUnit Experience
S01	JUnit TestNG	Novice (2-<=5 years)
S02	JUnit Nosetest	Novice (2-<=5 years)
S03	JUnit	Novice (2-<=5 years)
S04	JUnit, Testng, Mockito	Intermediate (5-<=10 years)
S05	JUnit, Mockito	Intermediate (5-<=10 years)
S06	Python Unittest	Limited experience (<2 years)
S07	Python Unittest	Limited experience (<2 years)

1.5 USO DE IDE

La calidad de la programación es, en principio, totalmente independiente del entorno de desarrollo utilizado (Un framework de testing puede hacer una diferencia, pero este aspecto se ha considerado en el apartado anterior). Sin embargo, la productividad puede variar en función de una multitud de aspectos (accesos directos vs menús, opciones de edición avanzadas, como la búsqueda y reemplazo en varios archivos, soporte para refactorización, etc.) En este experimento, Eclipse se utiliza como editor de texto. **Es de esperar que aquellos sujetos que usan IDEs similares pueden ser algo más productivo¹ que los sujetos que conozcan a otros entornos (por ejemplo: vi).**

La tabla 6 informa de los IDEs empleados regularmente por los sujetos experimentales. S01, S03, S04, S05 y S06 usan IDEs modernos como NetBeans, Eclipse, IntelliJ IDEA y Spider, que tienen aspecto similar. S02 utiliza un editor clásico como vim. S07 no informa. Ambos temas se destacan en la tabla 6.

Por lo tanto, el uso de IDE es una potencial variable moderadora. La eficacia de S02 y S07 puede disminuir debido a la falta de familiaridad con el IDE.

¹ Sin embargo, observe que la transición sin problemas entre los entornos de desarrollo modernos está lejos de ser clara. Cualquier persona habituada a cierto IDE puede informar de la frustración de la transición a una nueva, al menos por un corto período. La frustración es aún peor cuando los comandos típicos se encuentran en diferentes lugares sin un motivo claro. Por ejemplo, el comando Buscar y reemplazar puede estar ubicado en el menú Edición de una IDE, pero en el menú Refactorizar en otro. La transición de muy diferentes entornos de desarrollo, por ejemplo: vim puede ser más fácil ya que la interacción (atajos vs menús) es completamente diferente

Tabla 6. Uso de IDE

ID	IDEs used
S01	Netbeans
S02	VIM
S03	NetBeans
S04	Intellij Idea
S05	Eclipse
S06	Spyder Vim
S07	-

1.6 EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO

TDD es un enfoque de programación asociado a las prácticas ágiles. Las prácticas ágiles abogan por ciclos cortos de desarrollo y retroalimentación. Esas características también encajan en TDD. Entonces, es posible que la exposición anterior de prácticas ágiles favorezca el aprendizaje y aplicación de TDD. **Del mismo modo, los ciclos de vida tradicionales como la cascada, debido a su carácter secuencial, pueden interferir con el aprendizaje de TDD.**

Se preguntó a los sujetos acerca de su experiencia con metodologías de desarrollo. La Tabla 7 muestra los resultados. Todos los sujetos con la excepción de S04 han tenido una experiencia más o menos similar en desarrollo con enfoques tanto ágiles y cascada, por lo menos durante un año, aunque la media es de cuatro. S04 es la única excepción, reportando experiencia únicamente en desarrollos ágiles durante 10 años.

Como ninguno sujeto ha estado expuesto al ciclo de vida en cascada únicamente, **la experiencia con metodologías de desarrollo no puede moderar los resultados.**

Tabla 7. Metodologías de desarrollo

ID	Most recently used			Intermediate			Leastrecently used		
	Methd.	Educ. /yr	Usage /yr.	Methd.	Educ. /yr	Usage /yr.	Methd.	Educ. /yr	Usage /yr.
S01	Agile (Scrum)	0	5	Waterfall	0	4	-	-	-
S02	Agile (Scrum)	0	5	XP	0	2	Waterfall	0	2
S03	Agile (Scrum)	0	5	Iterative	0	10	Waterfall	3	10
S04	Agile(Scrum)	2	10	-	-	-	-	-	-
S05	Agile (Kanban)	0	1	Agile (Scrum)	0	1	Waterfall	1	0
S06	Agile (Scrum)	0	4	Waterfall	0	3	-	-	-
S07	Agile (Scrum)	0	3	Waterfall	0	6	-	-	-

1.7 TDD

TDD se relaciona con prácticas ágiles de programación. Esta variable es importante porque puede influir en el experimento. Si lo que queremos es comparar TDD con las metodologías tradicionales, el conocimiento que tengan sobre TDD de los sujetos que realizan el experimento podría ser una covariable.

La tabla 8. Muestra que los encuestados en su mayoría no están familiarizados con esta técnica. Sin embargo tres de ellos han usado alguna vez esta técnica. De los que han usado TDD no tienen experiencia o tienen muy poca como es el caso de S02.

Finalmente aunque sea poca la experiencia de los que han usado TDD, es necesario considerar la variable Experiencia en TDD como una covariable ya que dicho usuarios tienen cierta ventaja con respecto al resto.

Los resultados del análisis se muestran en la Tabla 8. Los usuarios (S01, S06 y S07) que no han usado TDD han dejado en blanco el apartado de la experiencia en TDD, se consideran por tanto sin experiencia: "No experience (0)". Después del análisis se decide que se separa los usuarios en dos categorías atendiendo la influencia sobre el experimento derivada de su experiencia en TDD: Los que influyen positivamente y los que influyen negativamente.

Los usuarios que no han usado TDD se considera que influyen negativamente, en esta categoría están: S01, S03, S06 y S07 (aparecen sombreados en la tabla 8). Y los que influyen positivamente son S02, S04 y S05. El usuario S02 ha usado TDD sin embargo no tiene experiencia, aun así se considera que tiene ventaja con respecto a los que no han usado TDD por tanto influye positivamente. El usuario S03 no ha usado TDD pero ha contestado que su experiencia es menor a 2 años, se entiende que se trata de un error de las opciones del formulario, quizá no estaba disponible la opción de

experiencia 0, y de ahí que los demás lo hayan dejado en blanco. Por tanto se considera que influye negativamente.

Tabla 8. Uso de TDD

ID	Usan TDD	Experiencia en TDD	Usan Katas	Experiencia usando Katas
S01	No	No experience (0)	No	No
S02	Yes	No experience (<2 years)	No	No
S03	No	No experience (<2 years)	No	No
S04	Yes	Novice (2-<=5 years)	No	No
S05	Yes	Novice (2-<=5 years)	No	No
S06	No	No experiencia (0)	No	No
S07	No	No experiencia (0)	No	No

Por otro lado el uso de Katas ayuda a los programadores a desarrollar y perfeccionar sus habilidades de programación a través de la práctica y la repetición. Se ha considerado que este factor puede influir a la hora de realizar el experimento. La tabla 8 muestra también los resultados de los formularios en este sentido. Como se puede ver, ningún usuarios usa Katas con lo cual esta variable no será tenida en cuenta como variable influyente.

1.8 RESUMEN DE LAS POSIBLES VARIABLES MODERADORAS (CO-VARIABLES)

En este apartado se resumen los resultados obtenidos en el tras el análisis. La tabla 9 muestra solo las variables que se consideran influyentes.

Un signo “+” indica que esa variable afecta de forma positiva y el signo “-” indica que puede influencias de forma negativa.

En análisis de los datos de los formularios proporciona información sobre otras variables que no se han considerado variables influyentes (covariables) estas son: Uso de Katas y Capacitación en TDD, puesto que los valores son los mismos para todos los sujetos que realizan el experimento.

1.9 CUANTIFICACIÓN DE LAS VARIABLES MODERADORAS

Con el objetivo de tener valores numéricos para poder usarlos como entrada de herramientas de análisis se elabora la tabla 10, contiene las variables moderadoras de la tabla 10 pero se ha proporcionado un valor numérico. A continuación se explica el criterio que se ha seguido para cuantificar cada variable:

1.9.1 CUALIFICACIÓN Y EXPERIENCIA

Esta variable se cuantifica teniendo en cuenta la preparación académica y la experiencia en la participación en proyectos software. Los valores de la tabla se corresponden con los años de experiencia de los individuos.

1.9.2 EXPERIENCIA DE PROGRAMACIÓN

Esta variable se cuantifica teniendo en cuenta sobre todo la experiencia que tienen los usuarios siendo desarrolladores. Los valores de la tabla se corresponden con los años de experiencia como desarrolladores usando el lenguaje de programación Java de cada uno de los usuarios. Se considera que lo importante es la experiencia que tienen los individuos en el desarrollo Java.

1.9.3 EXPERIENCIA EN TDD

Para representar la experiencia en TDD con valores numéricos se tiene en cuenta los años de experiencia usando TDD. Los que no han usado TDD tienen valor 0. Los que han usado TDD y no tienen experiencia tienen valor 1 y los que han usado TDD pero son novatos tienen valor 2.

1.9.4 EXPERIENCIA EN HERRAMIENTAS DE PRUEBAS DE UNIDAD Y JUNIT

Para cuantificar esta variable se tiene en cuenta la experiencia usando herramientas de pruebas de unidad y si además se trata de Junit. Los que no han usado Junit con valor 0. Los que son novatos con valor 2 y los intermedios con valor 5. Los valores se corresponden con los años de experiencia.

1.9.5 USO DE IDE

Para cuantificar esta variable se tiene en cuenta si están familiarizados con IDEs tipo Netbeans, Eclipse, VisualStudio, etc. Dado que no se puede cuantificar con un valor natural se va a establecer un valor booleano. De esta forma la traducción es: los usuarios que influyen positivamente le corresponde el valor 1 y a lo que influyen negativamente les corresponde el valor 0.

1.9.6 EXPERIENCIA EN PRUEBAS DE UNIDAD

Los valores de esta variable se corresponden con los años de experiencia de los usuarios realizando pruebas de unidad.

Tabla 9. Resumen de las posibles variables moderadoras

ID	Cualificación y Experiencia	Experiencia de programación	Experiencia en programación Java	Experiencia en Testing	Experiencia en frameworks de testing y JUnit	Uso de IDE	Experiencia en metodologías de desarrollo	Experiencia en TDD
S01	+	+	+	-	+	+	+	-
S02	+	+	+	+	+	-	+	+
S03	+	+	+	+	+	+	+	-
S04	+	+	+	+	+	+	+	+
S05	+	+	+	+	+	+	+	+
S06	-	-	-	-	-	+	+	-
S07	+	+	+	-	-	-	+	-

Tabla 10. Cuantificación de las variables moderadoras

ID	Cualificación y Experiencia	Experiencia de programación	Experiencia en programación Java	Experiencia en testing	Experiencia en frameworks de testing y JUnit	Uso de IDE	Experiencia en metodologías de desarrollo	Experiencia en TDD
S01	14	10	5	2	2	1	1	0
S02	6	10	2	5	2	0	1	1
S03	30	10	10	10	2	1	1	0
S04	10	10	5	5	5	1	1	2
S05	7.5	5	5	5	5	1	1	2
S06	0	2	0	2	0	1	1	0
S07	10	5	2	2	0	0	1	0

2 EJECUCIÓN FSECURE HELSINKI

Este apartado contiene un análisis de los resultados de las encuestas realizadas a los trabajadores de la empresa Helsinki.

2.1 CUALIFICACIÓN Y EXPERIENCIA DE LOS INDIVIDUOS

La información relativa a la cualificación y experiencia profesional de los participantes del experimento son las que se muestran en la tabla 11. Se puede apreciar que todos los individuos tienen carreras universitarias salvo FSH03 que todavía no la ha terminado, sin embargo es el que más experiencia tiene como desarrollador. Casi todos a excepción de FSH01, tienen carreras en ciencias de la computación. Todos, con excepción de FSH01 y FSH11 tienen experiencia como desarrolladores, experiencia que oscila entre los 3 y los 15 años.

El experimento compara metodologías de programación, con lo cual la experiencia que los individuos poseen como desarrolladores es un factor muy importante a tener en cuenta. En la tabla se muestra sombreado los individuos que se considera que están en desventaja con respecto al resto.

En el caso de FSH01 tiene dos aspectos a tener en cuenta, por un lado no tiene estudios de ciencias de la computación y por otro lado no tiene experiencia como desarrollador.

En el caso de FSH11, aunque es un individuo que posee muchas experiencia laboral en el ámbito de la TI, no tiene experiencia como programador, con lo cual se considera que también está en desventaja con respecto al resto de participantes.

Tabla 81. Cualificación y experiencia de los individuos

ID	Academic Degree	Current position	Professional career	Professional certifications
FSH01	MSc in physics	Quality Engineer	Tester: 5.5 years	Some testing
FSH02	MS Computer Science	Software Engineer	Software Developer: 15 years Product Manager: 1 year Team Manager: 1 year	Sun Certified Java Programmer (SCJP) Sun Certified Developer for Java Web Services (SCDJWS) Sun Certified Business Component Developer Java EE Platform 5 (SCBCD)

FSH03	Studied CS, but no graduation yet	CSO	Developer: 14 years	
FSH04	BS in Information Technology & Telecommunications	Labs	Testing 9 years Development ~9 years Management ~5 years i have many roles, but i write code 70% of the time	ISTQB testing certificate, however it's rather detached from real life
FSH05	BS in computer science, MS in computer science	Software developer	Developer: 3 years	
FSH11	BBA, computer science	R&D	Tester 4 years IT project manager 2 years Building security designer 2 years Technical support 1st level 2 years 2nd level 2 years 3rd level 2 years advanced level 2 years Grass cutting 6 months IT consultant 1 year	ITIL foundation MSCE in Exchange All F-Secure product certifications CSA, CSE (CheckPoint FW-1)
FSH13	Diplomi-insinööri, Computer science	RDO	Developer: 15 years	
FSH14	BA in Language Technology, MA in Language Technology (currently working on thesis)	Software Engineer	Developer 6 years, CS teacher/trainer 4 years, technical writer 3 months	

2.2 EXPERIENCIA DE PROGRAMACIÓN

La información sobre la experiencia de programación de los individuos es la que se muestra en la tabla 12. Se puede apreciar que hay tres individuos que no tienen experiencia en programación Java. Presentan por tanto una amenaza para el experimento dado que el experimento se lleva a cabo en este lenguaje. Los demás presentan poca experiencia en programación Java, esta oscila entre 0 y 6 años, salvo FSH02 que tiene 12 años de experiencia.

Tabla 12. Experiencia en el lenguaje de programación

ID	Most frequently used			Intermediate			Least frequently used		
	P.L.	Educ. /yr	Usage /yr.	P.L.	Educ. /yr	Usage /yr.	P.L.	Educ. /yr	Usage /yr.
FSH01	python	0	4						
FSH02	Java	0	12	Python	0	5	Bourne Shell	0	5
FSH03	Python	0	7	C	1	8	Java	1	0
FSH04	Python	0	9	Powershell	0	3	Perl, C++, C#, Misc	0	2
FSH05	Python	4	1	Java	3	2	JavaScript	1	2
FSH11	Python	0	3	Shell scripting	0	6			
FSH13	c	1	0	c++	1	9	java	1	6
FSH14	Python	2	6	Java	3	4	C	1	3

La tabla 13 muestra la experiencia en programación y en particular en programación Java que proporcionaron los individuos en la encuesta. Hay consistencia entre las tablas 12 y 13 ya que los que se consideraban una amenaza para el experimento asumen no tener experiencia en programación Java (FSH01, FSH04 y FSH11).

Tabla 13. Programación y experiencia en Java

ID	Programming experience	Java experience
FSH01	Novice (2-<=5 years)	No experience (<2 years)
FSH02	Expert (>10 years)	Expert (>10 years)
FSH03	Intermediate (5-<=10 years)	Novice (2-<=5 years)
FSH04	Intermediate (5-<=10 years)	No experience (<2 years)
FSH05	Intermediate (5-<=10 years)	Novice (2-<=5 years)
FSH11	Expert (>10 years)	No experience (<2 years)
FSH13	Expert (>10 years)	Intermediate (5-<=10 years)
FSH14	Expert (>10 years)	Intermediate (5-<=10 years)

2.3 EXPERIENCIA EN TESTING

La información de la tabla 14 presenta la experiencia de los individuos realizando pruebas de unidad. La mayoría de los individuos poseen poca experiencia salvo FSH03 que es intermedio y FSH02 que se presenta como experto, información que concuerda con su experiencia profesional. Sin embargo los individuos FSH01 y FSH11 No tienen experiencia. Con estos datos podemos determinar que los individuos que no tienen

ninguna experiencia se encuentran en desventaja con respecto al resto y por consiguiente influyen negativamente en el experimento.

Tabla 14. Experiencia en Unit testing

ID	Experiencia en Unit testing
FSH01	No experience (<2 years)
FSH02	Expert (>10 years)
FSH03	Intermediate (5-<=10 years)
FSH04	Novice (2-<=5 years)
FSH05	Novice (2-<=5 years)
FSH11	No experience (<2 years)
FSH13	Novice (2-<=5 years)
FSH14	Novice (2-<=5 years)

2.4 EXPERIENCIA EN FRAMEWORKS DE TESTING

Dado que la aplicación de TDD requiere una infraestructura de pruebas compuestas por determinadas herramientas, se considera que poseer experiencia previa en estas herramientas puede influir positivamente en el experimento. En concreto la experiencia en JUnit es muy influyente en el experimento puesto que las pruebas son parte importante del ciclo de TDD.

La información relativa a la experiencia usando herramientas y frameworks de testing por parte de los individuos se recoge en la tabla 15. Se puede apreciar que en general, salvo FSH11 todos han usado herramientas de pruebas de unidad. Sin embargo, y en concordancia con la tabla 12 que recoge la experiencia de programación, estas herramientas están relacionadas con el lenguaje de programación Python. En general los individuos tienen poca experiencia con JUnit que es el framework para pruebas de unidad para el lenguaje Java, solo FSH02 es un experto.

Hay que destacar también que la tabla 15 es consistente solo para los individuos FSH02 y FSH13, ya que mencionan JUnit en las herramientas de testeo utilizadas. FSH05 y FSH14 dicen ser Novato e intermedio respectivamente sin embargo no mencionaron a JUnit en las herramientas de testeo. Habrá que ver hasta qué punto esto es verdad.

Tabla 15. Unit testing tools and JUnit experience

ID	Unit test tools used	JUnit Experience
FSH01	python nose	No experience (<2 years)
FSH02	JUnit, Mockito, Jenkins	Expert (>10 years)
FSH03	nosetests	No experience (<2 years)
FSH04	Nose tests for python	No experience (<2 years)
FSH05	nosetests, jasmine, karma, Test::More	Novice (2-<=5 years)
FSH11	No	No experience (<2 years)
FSH13	junit, nosetest, mockito	Novice (2-<=5 years)
FSH14	Nosier, py.test, unittest (Python standard library)	Intermediate (5-<=10 years)

2.5 USO DE IDES

El IDE que se usa en el experimento es Eclipse, se considera de mucha relevancia conocer el entorno de desarrollo puesto que puede facilitar mucho el trabajo de codificación ya que proporciona muchas utilidades que ahorran tiempo durante la programación.

La tabla 16 recoge la información relativa al uso de IDEs por parte de los participantes del experimento. El único que presenta experiencia usando Eclipse es FSH11. Se considera positivo el conocimiento en IntelliJ IDEA al ser un entorno de desarrollo para Java. En el caso de FSH04 se considera que aunque no ha trabajado con un IDE para el lenguaje Java, la experiencia de manejar Komodo IDE puede influir positivamente frente a los usuarios que no hay usado IDEs (FSH01, FSH03, FSH05 y FSH14).

El resto de participantes (FSH01, FSH03, FSH05 y FSH14) no tienen experiencia usando IDEs con lo cual se considera que esto puede influir de forma negativa en el experimento.

Tabla 16. Uso de IDEs

ID	IDEs used
FSH01	basic text editor i.e. no IDE
FSH02	IntelliJ IDEA
FSH03	vi
FSH04	Komodo IDE
FSH05	SublimeText
FSH11	Eclipse
FSH13	intellij idea, pycharm
FSH14	None (Sublime text + command line)

2.6 EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO

La tabla 17 muestra la experiencia de los individuos en metodologías de desarrollo. Todos los individuos presentan experiencia en metodologías ágiles, sin embargo FSH01, FSH05 están en desventaja con respecto al resto de sujetos dado que solo tienen experiencia en el modelo de desarrollo ágil, con lo cual se considera puede influir en el experimento.

Tabla 17. Metodologías

ID	Most recently used			Intermediate			Leastrecently used		
	Methd.	Edu c. /yr	Usag e /yr.	Methd.	Edu c. /yr	Usag e /yr.	Methd.	Edu c. /yr	Usag e /yr.
FSH01	Agile	0	5						
FSH02	Waterfall	1	3	Iterative	0	7	Scrum	0	5
FSH03	agile/scrum	0	7	agile/kanban	0	4	waterfall	1	6
FSH04	Waterfall	0	3	Scrum	0	3	Kanban	0	3
FSH05	agile, scrum	0	3						
FSH11	Agile, Scrum	0	2	Agile, Kanban	0	2			
FSH13	waterfall/iterative	0	10	scrum	0	4	kanban	0	1
FSH14	Scrum	1	2	tdd	1	0	waterfall	0	6

2.7 EXPERIENCIA EN TDD

La experiencia en TDD es importante porque puede influir en el experimento. Si lo que queremos es comparar TDD con las metodologías tradicionales, el conocimiento que tengan sobre TDD de los sujetos que realizan el experimento podría ser una covariable.

La tabla 18 muestra que los encuestados en su mayoría no están familiarizados con esta técnica. Los que han usado TDD han contestado que no tienen experiencia. Aquí hay que puntualizar que la categoría “No experiencia (< 2 years)” tiene un intervalo muy grande y no permite diferenciar la experiencia entre los individuos. Sin embargo sí que se pueden diferenciar dos categorías. Los que han usado TDD y los que No. De esta forma podemos afirmar que los usuarios que no han usado TDD (FSH02, FSH11 y FSH13) influyen negativamente.

Aunque sea poca la experiencia de los que han usado TDD, es necesario considerar la variable Experiencia en TDD como una covariable ya que dicho usuarios tienen cierta ventaja con respecto al resto.

Table 18. Uso de TDD

ID	Usan TDD	Experiencia en TDD	Training on TDD	Usan Katas	Experiencia usando Katas
FSH01	Yes	No experience (<2 years)	No	No	
FSH02	No	No experience (<2 years)	No	No	
FSH03	Yes	Novice (2-<=5 years)	Yes	No	
FSH04	Yes	No experience (<2 years)	No	No	
FSH05	Yes	No experience (<2 years)	No	Yes	Can't remember, there were only a few
FSH11	No		No	No	
FSH13	No		No	Yes	Ericsson session
FSH14	Yes	No experience (<2 years)	Yes	No	

El uso de Katas ayuda a los programadores a desarrollar y perfeccionar sus habilidades de programación a través de la práctica y la repetición. Se ha considerado que este factor puede influir a la hora de realizar el experimento. La tabla 20 muestra la información relativa al uso de katatas por parte de los individuos. FSH05 y FSH13 han usado Katatas y se considera que esto influye positivamente en el experimento.

2.8 RESUMEN DE LAS POSIBLES VARIABLES MODERADORAS (CO-VARIABLES)

En este apartado se resumen los resultados obtenidos en el tras el análisis. La tabla 21 muestra solo las variables que se consideran influyentes.

Un signo “+” indica que esa variable afecta de forma positiva y el signo “-” indica que puede influencias de forma negativa.

2.9 CUANTIFICACIÓN DE LAS VARIABLES MODERADORAS

Con el objetivo de tener valores numéricos para poder usarlos como entrada de herramientas de análisis se elabora la tabla 22. Contiene las variables moderadoras de la tabla 21 pero se ha proporcionado un valor numérico. A continuación se explica el criterio que se ha seguido para cuantificar cada variable:

2.9.1 CUALIFICACIÓN Y EXPERIENCIA

Esta variable se cuantifica teniendo en cuenta la preparación académica y la experiencia en la participación en proyectos software. Los valores de la tabla se corresponden con los años de experiencia de los individuos.

2.9.2 EXPERIENCIA DE PROGRAMACIÓN

Esta variable se cuantifica teniendo en cuenta sobre todo la experiencia que tienen los usuarios siendo desarrolladores. Los valores de la tabla se corresponden con los años de experiencia como desarrolladores usando el lenguaje de programación Java de cada uno de los usuarios. Se considera que lo importante es la experiencia que tienen los individuos en el desarrollo Java.

2.9.3 EXPERIENCIA DE PROGRAMACIÓN EN JAVA

Los valores de la tabla se corresponden con los años de experiencia como desarrolladores usando el lenguaje de programación Java de cada uno de los usuarios. Se considera que lo importante es la experiencia que tienen los individuos en el desarrollo Java.

2.9.4 EXPERIENCIA EN PRUEBAS DE UNIDAD

Los valores de esta variable se corresponden con los años de experiencia de los usuarios realizando pruebas de unidad.

2.9.5 EXPERIENCIA EN HERRAMIENTAS DE PRUEBAS DE UNIDAD Y JUNIT

Para cuantificar esta variable se tiene en cuenta la experiencia usando herramientas de pruebas de unidad y si además se trata de Junit. Los que no han usado Junit con valor 0. Los que son novatos con valor 2, los intermedios con valor 5 y expertos con valor 10. Los valores se corresponden con los años de experiencia.

2.9.6 USO DE IDE

Para cuantificar esta variable se tiene en cuenta si están familiarizados con IDEs tipo Netbeans, Eclipse, VisualStudio, etc. Dado que no se puede cuantificar con un valor

natural se va a establecer un valor booleano. De esta forma la traducción es: los usuarios que influyen positivamente le corresponde el valor 1 y a lo que influyen negativamente les corresponde el valor 0.

2.9.7 EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO

Para cuantificar esta variable se ha tenido en cuenta en número de años que los sujetos tienen como experiencia en desarrollos ágiles.

2.9.8 EXPERIENCIA EN TDD

Para representar la experiencia en TDD con valores numéricos se tiene en cuenta dos cosas: Si han usado TDD y los años de experiencia usando TDD. Los que no han usado TDD tienen valor 0. Los que han usado TDD y no tienen experiencia tienen valor 1 y los que han usado TDD pero son novatos tienen valor 2.

Tabla 19. Resumen de las posibles variables moderadoras

ID	Cualificación y Experiencia	Experiencia de programación	Experiencia en programación Java	Experiencia en testing	Experiencia en frameworks de testing y JUnit	Uso de IDE	Experiencia en metodologías de desarrollo	Experiencia en TDD
FSH01	-	-	-	-	-	-	-	+
FSH02	+	+	+	+	+	+	+	-
FSH03	+	+	+	+	-	-	+	+
FSH04	+	+	-	+	-	+	+	+
FSH05	+	+	+	+	+	-	-	+
FSH11	-	+	-	-	-	+	+	-
FSH13	+	+	+	+	+	+	+	-
FSH14	+	+	+	+	+	-	+	+

Tabla 20. Cuantificación de las variables moderadoras

ID	Cualificación y Experiencia	Experiencia de programación	Experiencia en programación Java	Experiencia en Testing	Experiencia en herramientas de pruebas de unidad y JUnit	Uso de IDE	Experiencia en metodologías de desarrollo	Experiencia en TDD
FSH01	0	2	0	0	0	0	0	1
FSH02	15	10	10	10	10	1	1	0
FSH03	14	5	2	5	0	0	1	2
FSH04	9	5	0	2	0	1	1	1
FSH05	3	5	2	2	2	0	0	1
FSH11	0	10	0	0	0	1	1	0
FSH13	15	10	5	2	2	1	1	0
FSH14	6	10	5	2	5	0	1	1

3 EJECUCIÓN FSECURE KUALA LUMPUR

Este apartado contiene un análisis de los datos demográficos reportados por los sujetos de la empresa Kuala Lumpur.

3.1 CUALIFICACIÓN Y EXPERIENCIA DE LOS INDIVIDUOS

La información relativa a la cualificación y experiencia profesional de los participantes del experimento son las que se muestran en la tabla 21. Se puede apreciar que todos los individuos tienen carreras universitarias. En su mayoría en Ciencias de la computación. Hay algunos que tiene ingeniería eléctrica (FSK02, FSK04, FSK11 y FSK14), sin embargo su vida profesional la desarrollan como ingenieros de software.

El experimento compara metodologías de programación, con lo cual la experiencia que los individuos poseen como desarrolladores es un factor muy importante a tener en cuenta. En la tabla 21 se muestra sombreado los individuos que se considera que están en desventaja con respecto al resto.

Los individuos FSK03, FSK06 y FSK19 están en desventaja con respecto al resto puesto que no tienen experiencia como desarrolladores y se considera que esto puede influir negativamente en el experimento.

Tabla 21. Cualificación y experiencia de los individuos

ID	Academic Degree	Current position	Professional career	Professional certifications
FSK01	Certificate in Electronic Eng, Certificate in Computing, Diploma in Computing BS in Software Engineering	Engineer	Developer 6 years	
FSK02	BEng in Electronics Engineering (major Computer)	Software Engineer	Software Engineer 6+ years	
FSK03	Bachelor in Information Technology	Software Engineer	Quality Engineer(Intern) (6 months) Software Engineer (3 months)	CEH (Certified Ethical Hacker)
FSK04	B.Eng Electrical Electronic	Software Engineer	R&D Software Engineer --> 9 years	
FSK05	BS in	Software Engineer	Software Engineer - 5 years	SCJP1.4

	Information Technology			
FSK06	BS in computer science	r&d	tester - 2 years	
FSK07	Master's degree in Networking and Computing	R&D	Software Engineer 5 years	
FSK08	BS in Microelectronic	Developer	developer 10years	
FSK09	Snr. Software Engineer Bsc(hons) Computing, University of Bolton, UK	RnD	developer 5 years tester 1 year	
FSK11	BS in Computer Engineering	Software Engineer	Web Developer for 5 years	
FSK12	BS in computer science	Developer	Developer 11 years	-
FSK13	ms in computer science	r&d	developer 5 years tester 1 year	Sun java certs MCSE
FSK14	B.Eng in Electronics Engineering.	SENIOR SOFTWARE ENGINEER	Developer: 9 years.	
FSK16	BS in computer Science	Software Engineer	Software Developer 15 years	
FSK18	BS in Information System Engineering major computer science	Software Engineer	Software Engineer 81	MS SQL Server 2005, Angular JS
FSK19	BS electrical n electronic engineering majoring in Computer	TDD	Tester 7 years	ISTQB

3.2 EXPERIENCIA DE PROGRAMACIÓN

La información sobre la experiencia de programación de los individuos es la que se muestra en la tabla 22. Hay 5 individuos que no tienen experiencia en programación Java. Presentan por tanto una amenaza para el experimento dado que el experimento se lleva a cabo en este lenguaje.

Tabla 92. Experiencia en el lenguaje de programación

ID	Most frequently used			Intermediate			Least frequently used		
	P.L.	Educ. /yr	Usage /yr.	P.L.	Educ. /yr	Usage /yr.	P.L.	Educ. /yr	Usage /yr.
FSK01	Java	3	0	PHP	0	8	Javascript	0	10
FSK02	C/C++	1	4	Javascript	0	2	Python	0	1
FSK03	Java	4.5	<1 year	PHP	2	0	C	1	0
FSK04	C++	1	7	Python	0	2.5	C#	0	4
FSK05	Java	0.5	2	C++	0.5	1	Python	0	1
FSK06	python	0	2	java	1	0	C++	1	0
FSK07	Java	3	3.5	Python	0	2	C/C++	5	3
FSK08	C/C++	4	7	Java	2	3	python	0	1
FSK09	C++	8	8	Python	3	3	Javascript	3	3
FSK11	PHP	1	3	Javascript	1	5	Python	0	1
FSK12	Java	4	11	Scala	0	2	Javascript	1	7
FSK13	c++	2	2	java	2	2	python	0	1
FSK14	C	3	3	Python	0	7	Javascript	0	2
FSK16	PHP	0	12	Java	1	2	JavaScript	0	10
FSK18	PHP	0	3	C#	0.4	1	Java	0.4	0.8
FSK19	Python	0	2	Java	1	2	C#	0	2

La tabla 22 contiene la información relativa a la experiencia de programación de los individuos. En concreto nos interesa la experiencia de programación en el lenguaje Java. Esta información se encuentra en la tabla 23. Se considera que un conocimiento previo del lenguaje en el que se realiza el experimento puede influir positivamente. La mayoría de los individuos tiene cierto conocimiento del lenguaje Java. Sin embargo hay tres individuos (FSK02, FSK11 y FSK14) que no tienen experiencia en dicho lenguaje. Se considera que esto puede influir negativamente en el experimento. El usuario FSK01 no tiene experiencia sin embargo ha estudiado 3 años el lenguaje Java. Y FSK06 un año. FSK18 también pone que no tienen experiencia aunque en la tabla anterior se puede apreciar que ha estudiado y además tiene casi un año de experiencia en Java. Idem para el usuario FSK18.

Tabla 23. Programación y experiencia en Java

ID	Programming experience	Java experience
FSK01	Intermediate (5-<=10 years)	No experience (<2 years)
FSK02	Intermediate (5-<=10 years)	No experience (<2 years)
FSK03	Novice (2-<=5 years)	Novice (2-<=5 years)
FSK04	Intermediate (5-<=10 years)	Novice (2-<=5 years)
FSK05	Intermediate (5-<=10 years)	Novice (2-<=5 years)
FSK06	Novice (2-<=5 years)	No experience (<2 years)
FSK07	Intermediate (5-<=10 years)	Novice (2-<=5 years)
FSK08	Intermediate (5-<=10 years)	Novice (2-<=5 years)
FSK09	Intermediate (5-<=10 years)	Novice (2-<=5 years)
FSK11	Intermediate (5-<=10 years)	No experience (<2 years)
FSK12	Expert (>10 years)	Expert (>10 years)
FSK13	Intermediate (5-<=10 years)	Novice (2-<=5 years)
FSK14	Intermediate (5-<=10 years)	No experience (<2 years)
FSK16	Expert (>10 years)	Novice (2-<=5 years)
FSK18	Novice (2-<=5 years)	No experience (<2 years)
FSK19	Intermediate (5-<=10 years)	Novice (2-<=5 years)

3.3 EXPERIENCIA EN TESTING

La información de la tabla 24 presenta la experiencia de los individuos realizando pruebas de unidad. La mayoría de los individuos poseen poca experiencia. Con estos datos podemos determinar que los individuos que no tienen ninguna experiencia se encuentran en desventaja con respecto al resto y por consiguiente influyen negativamente en el experimento.

Tabla 24. Experiencia en Unit testing

ID	Experiencia en Unit testing
FSK01	Novice (2-<=5 years)
FSK02	Novice (2-<=5 years)
FSK03	No experience (<2 years)
FSK04	Novice (2-<=5 years)
FSK05	Novice (2-<=5 years)
FSK06	No experience (<2 years)
FSK07	Novice (2-<=5 years)
FSK08	No experience (<2 years)
FSK09	Novice (2-<=5 years)
FSK11	No experience (<2 years)
FSK12	Novice (2-<=5 years)
FSK13	Novice (2-<=5 years)
FSK14	Novice (2-<=5 years)
FSK16	No experience (<2 years)
FSK18	Novice (2-<=5 years)
FSK19	No experience (<2 years)

3.4 EXPERIENCIA EN FRAMEWORKS DE TESTING

La tabla 25 contiene la información relativa a la experiencia de los individuos usando herramientas de pruebas de unidad. En concreto nos interesa saber si han usado JUnit. Se considera que influye positivamente que los individuos tengan cierto conocimiento de JUnit. Los que no tienen experiencia y por tanto influyen negativamente aparecen sombreados en la tabla 25.

Tabla 25. Unit testing tools and JUnit experience

ID	Unit test tools used	JUnit Experience
FSK01	Depends on the project. Current favorite tool is: Karma	No experience (<2 years)
FSK02	C++ - IBM Test RealTime, GoogleMock Javascript - Jasmine, JSTestDriver Python - nose, mock, Coverage	No experience (<2 years)
FSK03	JUnit/Mockito	No experience (<2 years)
FSK04	Jasmin for Python	Novice (2-<=5 years)
FSK05	Jasmine JUnit	Novice (2-<=5 years)
FSK06	no	No experience (<2 years)
FSK07	django-test, jenkins	No experience (<2 years)
FSK08	junit, jasmine	No experience (<2 years)
FSK09	No	Novice (2-<=5 years)
FSK11	Karma	No experience (<2 years)
FSK12	TestNG	Novice (2-<=5 years)
FSK13	no	Novice (2-<=5 years)
FSK14	Karma nosetests python unittest	No experience (<2 years)
FSK16	Django Unit Test Jasmine Javascript Unit test	No experience (<2 years)
FSK18	JUnit	No experience (<2 years)
FSK19	Not doing unit testing	No experience (<2 years)

3.5 USO DE IDES

El IDE que se usa en el experimento es Eclipse, se considera de mucha relevancia conocer el entorno de desarrollo puesto que puede facilitar mucho el trabajo de codificación ya que proporciona muchas utilidades que ahorran tiempo durante la programación.

La tabla 26 recoge la información relativa al uso de IDEs por parte de los participantes del experimento. Hay tres individuos que no han usado un IDE (FSK13, FSK14 y FSK19) y se considera que influyen negativamente en el experimento. El resto aunque algunos

no han usado Eclipse, si que tienen experiencia usando IDEs que funcionan de forma similar.

Tabla 26. Uso de IDEs

ID	IDEs used
FSK01	IntelliJ IDEA
FSK02	NetBeans
FSK03	Eclipse
FSK04	PyCharm
FSK05	PyStorm
FSK06	eclipse with PyDev
FSK07	pycharm. eclipse
FSK08	eclipse, phpstorm
FSK09	MSVS studio
FSK11	PyCharm, Php Storm
FSK12	IntelliJ IDEA
FSK13	vi sublime text
FSK14	Pyscripter, Geany
FSK16	Pycharm
FSK18	Eclipse
FSK19	Not doing Unit Testing

3.6 EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO

La tabla 27 muestra la experiencia de los individuos en metodologías de desarrollo. Entendemos que la experiencia en metodologías ágiles es un aspecto positivo a tener en cuenta. Sin embargo, Los sujetos que aparecen sombreados en la tabla 27 están en desventaja con respecto al resto teniendo en cuenta que solo se han enfrentado a una de las metodologías. Por tanto la experiencia con metodologías de desarrollo puede ser una variable influyente o moderadora.

Tabla 27. Metodologías

ID	Most recently used			Intermediate			Least recently used		
	Methd.	Educ. /yr	Usage /yr.	Methd.	Educ. /yr	Usage /yr.	Methd.	Educ. /yr	Usage /yr.
FSK01	Agile (scrum)	0	4	Waterfall	1	1			
FSK02	waterfall	0	3.5	scrum	0	3			
FSK03	Scrum	1	<1 year	RUP	1	0			
FSK04	agile - scrum	0	3						
FSK05	scrum	0	3	waterfall	3	1	iterative	0	1
FSK06	agile (scrum)	0	2	waterfall	3	0			
FSK07	Scrum	0	5						
FSK08	waterfall	4	7	agile	0	2			
FSK09	Agile (scrum)	1	5						
FSK11	agile scrum	0	3						
FSK12	waterfall	4	6	Scrum	0	5			
FSK13	waterfall	2	4	Agile scrum	1	5			
FSK14	Agile	0	3	Scrum	0	3			
FSK16	Scrum	0	1	Waterfall	2	10	Iterative	2	2
FSK18	agile - scrum	0	1.8	waterfall	0.4	4			
FSK19	Agile	0	2	V Model	0	3			

3.7 EXPERIENCIA EN TDD

La experiencia en TDD es importante porque puede influir en el experimento. Si lo que queremos es comparar TDD con las metodologías tradicionales, el conocimiento que tengan sobre TDD de los sujetos que realizan el experimento podría ser una covariable.

La tabla 28. Muestra que los encuestados en su mayoría no están familiarizados con esta técnica. En la tabla 28 se muestran sombreados los individuos que se considera que influyen negativamente en el experimento puesto que no tienen experiencia en TDD. Recaltar que el individuo FSK09 aunque dice que no ha usado TDD se considera novato. Se entiende que se trata de un error suyo al rellenar el formulario.

Table 28. Uso de TDD

ID	Usan TDD	Experiencia en TDD	Usan Katas	Experiencia usando Katas
FSK01	Yes	No experience (<2 years)	Yes	Custom generated tasks
FSK02	No		No	
FSK03	No	No experience (<2 years)	No	
FSK04	Yes	Novice (2-<=5 years)	No	
FSK05	No	No experience (<2 years)	Yes	We conduct this KATA internally in own team last year 2013. Doing "Simply Buzz", "Bowling" task.
FSK06	No		Yes	calculate the bowling marks
FSK07	Yes	No experience (<2 years)	Yes	FizzBuzz RomanNumerals
FSK08	Yes	No experience (<2 years)	Yes	bowling
FSK09	No	Novice (2-<=5 years)	No	
FSK11	No	No experience (<2 years)	Yes	Online: http://codingdojo.org/cgi-bin/wiki.pl?KataFizzBuzz
FSK12	Yes	Novice (2-<=5 years)	No	
FSK13	No		No	
FSK14	No	No experience (<2 years)	No	
FSK16	No	No experience (<2 years)	No	
FSK18	Yes	No experience (<2 years)	No	
FSK19	No	No experience (<2 years)	No	

El uso de Katas ayuda a los programadores a desarrollar y perfeccionar sus habilidades de programación a través de la práctica y la repetición. Se ha considerado que este factor puede influir a la hora de realizar el experimento.

La tabla 28 muestra la información relativa al uso de katas por parte de los individuos. La mayoría de los individuos no está familiarizada con el uso de Katas, sin embargo hay un número considerable que sí. Los que no están familiarizados con el uso de Katas aparecen sombreados en la tabla 28 y se considera que influyen negativamente en el experimento.

3.8 RESUMEN DE LAS POSIBLES VARIABLES MODERADORAS (CO-VARIABLES)

En este apartado se resumen los resultados obtenidos en el tras el análisis. La tabla 32 muestra las variables que se consideran influyentes.

Un signo “+” indica que esa variable afecta de forma positiva y el signo “-” indica que puede influencias de forma negativa.

3.9 CUANTIFICACIÓN DE LAS VARIABLES MODERADORAS

Con el objetivo de tener valores numéricos para poder usarlos como entrada de herramientas de análisis se elabora la tabla 11, contiene las variables moderadoras de la tabla 10 pero se ha proporcionado un valor numérico. A continuación se explica el criterio que se ha seguido para cuantificar cada variable:

3.9.1 CUALIFICACIÓN Y EXPERIENCIA

Esta variable se cuantifica teniendo en cuenta la preparación académica y la experiencia en la participación en proyectos software. Los valores de la tabla se corresponden con los años de experiencia de los individuos.

3.9.2 EXPERIENCIA DE PROGRAMACIÓN

Para cuantificar esta variable se tiene en cuenta directamente los años de experiencia que han reportado los sujetos.

3.9.3 EXPERIENCIA DE PROGRAMACIÓN JAVA

Esta variable se cuantifica teniendo en cuenta sobre todo la experiencia que tienen los usuarios siendo desarrolladores. Los valores de la tabla se corresponden con los años de experiencia como desarrolladores usando el lenguaje de programación Java de cada uno de los usuarios. Se considera que lo importante es la experiencia que tienen los individuos en el desarrollo Java. Para que la información de tabla con información numérica se lo más fiel con la realidad se ha decidido poner el valor “1” para el caso en el que los individuos han usado Java y sin embargo han indicado no tener experiencia. El resto de valores se corresponde con los años de experiencia de cada individuo.

3.9.4 EXPERIENCIA EN PRUEBAS DE UNIDAD

Los valores de esta variable se corresponden con los años de experiencia de los usuarios realizando pruebas de unidad

3.9.5 EXPERIENCIA EN HERRAMIENTAS DE PRUEBAS DE UNIDAD Y JUNIT

Para cuantificar esta variable se tiene en cuenta la experiencia usando herramientas de pruebas de unidad y si además se trata de JUnit. Los que no han usado JUnit con valor 0. Los que son novatos con valor 2 y los intermedios con valor 5. Los valores se corresponden con los años de experiencia.

3.9.6 USO DE IDE

Para cuantificar esta variable se tiene en cuenta si están familiarizados con IDEs tipo Netbeans, Eclipse, VisualStudio, etc. Dado que no se puede cuantificar con un valor natural se va a establecer un valor booleano. De esta forma la traducción es: los usuarios que influyen positivamente le corresponde el valor 1 y a lo que influyen negativamente les corresponde el valor 0.

3.9.7 METODOLOGÍAS DE DESARROLLO

Esta variable se cuantifica con el número de años de experiencia que reportan los sujetos en metodologías de desarrollo.

3.9.8 EXPERIENCIA EN TDD

Para representar la experiencia en TDD con valores numéricos se tiene en cuenta los años de experiencia usando TDD. Los que no han usado TDD tienen valor 0. Los que han usado TDD y no tienen experiencia tienen valor 1 y los que han usado TDD pero son novatos tienen valor 2.

Tabla 29. Resumen de las posibles variables moderadoras

ID	Culificaci ón y Experiencia	Experiencia de programación	Experiencia de programación en Java	Experiencia en pruebas de unidad	Experiencia en herramientas de pruebas de unidad y Junit	Us o de IDE	Experiencia en metodologías de desarrollo	Experiencia en TDD	Uso de Katas
FSK01	+	+	-	+	-	+	+	+	+
FSK02	+	+	-	+	-	+	+	-	-
FSK03	-	-	+	-	+	+	-	-	-
FSK04	+	+	+	+	+	+	+	+	-
FSK05	+	+	+	+	+	+	+	-	+
FSK06	-	-	-	-	-	+	-	-	+
FSK07	+	+	+	+	-	+	-	+	+
FSK08	+	+	+	-	+	+	+	+	+
FSK09	+	+	+	+	+	+	+	+	-
FSK11	+	+	-	-	-	+	-	-	+
FSK12	+	+	+	+	+	+	+	+	-
FSK13	+	+	+	+	+	-	+	-	-
FSK14	+	+	-	+	-	-	+	-	-
FSK16	+	+	+	-	-	+	+	-	-
FSK18	+	-	-	+	+	+	+	+	-
FSK19	-	+	+	-	-	-	+	-	-

Tabla 30. Cuantificación de las variables moderadoras

ID	Culificaci ón y Experiencia	Experiencia de programación	Experiencia de programación en Java	Experiencia en pruebas de unidad	Experiencia en herramientas de pruebas de unidad y Junit	Us o de IDE	Experiencia en metodologías de desarrollo	Experiencia en TDD	Uso de Katas
FSK01	6	5	0	2	0	1	1	1	1
FSK02	6	5	0	2	0	1	1	0	0
FSK03	0.25	2	2	0	0	1	0	0	0
FSK04	9	5	2	2	2	1	1	2	0
FSK05	5	5	2	2	2	1	1	0	1
FSK06	0	2	0	0	0	1	0	0	1
FSK07	5	5	2	2	0	1	0	1	1
FSK08	10	5	2	0	0	1	1	1	1
FSK09	5	5	2	2	2	1	1	2	0
FSK11	5	5	0	0	0	1	0	0	1
FSK12	11	10	10	2	2	1	1	2	0
FSK13	5	5	2	2	2	0	1	0	0
FSK14	9	5	0	2	0	0	1	0	0
FSK16	15	10	2	0	0	1	1	0	0
FSK18	9	2	0	2	0	1	1	1	0
FSK19	0	5	2	0	0	0	1	0	0

4 EJECUCIÓN ELEKTROBIT KAJAANI

Este apartado contiene un análisis de los resultados de las encuestas realizadas a los trabajadores de la empresa Electrobit.

4.1 CUALIFICACIÓN Y EXPERIENCIA DE LOS INDIVIDUOS

La información relativa a la cualificación y experiencia profesional de los participantes del experimento son las que se muestran en la tabla 31. Se puede apreciar que todos los individuos tienen carreras universitarias. En su mayoría en Ciencias de la computación.

El experimento compara metodologías de programación, con lo cual la experiencia que los individuos poseen como desarrolladores es un factor muy importante a tener en cuenta. En la tabla 31 se muestra sombreado los individuos que se considera que están en desventaja con respecto al resto.

La mayoría de los individuos tienen experiencia como desarrolladores

Los individuos ebkaj18 y ebkaj21 están en desventaja con respecto al resto puesto que no tienen experiencia como desarrolladores y se considera que esto puede influir negativamente en el experimento.

Tabla31. Cualificación y experiencia de los individuos

ID	Academic Degree	Current position	Professional career	Professional certifications
ebkaj11	Degree in information processing (Kajaani polytechnic)	Senior SW designer	system testing 6 months. functional testing 2 years developer 8 years.	scrum master certification
ebkaj13	Bachelor's Degree in computer science (Embedded SW)	Senior Designer	Developer 2 years SW Team leader 2 years SW Maintenance Technical specialist 2 years SW Specification team member 3 years SW Architect 6 months	
ebkaj14	Undergraduate	Software Engineer	developer 13 years	
ebkaj15	BS in computer science	Senior software engineer	SW developer 7 years, SW tester 2 years, build manager 2 years	
ebkaj16	BSc in computer science	Design Engineer	test/integrator 1 year build manager/integrator 5 years developer 4 years ui designer/usability 0,5	

			years	
ebkaj17	BS	Specialist	Developer 15 years, Specification tasks 5 years,	MCP (Microsoft Certified Professional) 1997 using FoxBase
ebkaj18	Bachelor Degree of Science, Information Technology	Tester	Tester 1 year 5 month	
ebkaj20	BS in Electrical Engineering	Specialist	developer 20 years	
ebkaj21	BS in electronics engineering.	SW dev	SW designer 1 year, Scrum Master 13 months	

4.2 EXPERIENCIA DE PROGRAMACIÓN

La información sobre la experiencia de programación de los individuos es la que se muestra en la tabla 32. Todos los individuos presentan experiencia con el lenguaje C++. Por lo tanto la experiencia de programación de momento no influye en el experimento.

Tabla 32. Experiencia en el lenguaje de programación

ID	Most frequently used			Intermediate			Least frequently used		
	P.L.	Educ /yr	Usag e /yr.	P.L.	Educ /yr	Usag e /yr.	P.L.	Educ /yr	Usag e /yr.
ebkaj1 1	C++	1	7	c	2	1	Java	3	0
ebkaj1 3	C++	1	4	C	1	4			
ebkaj1 4	c++	0 years	13						
ebkaj1 5	c	2	5	C++	2	7	pytho n	2	1
ebkaj1 6	C/C++	0,5	5	perl/php/scriptin g	0	4	java	0	0,5
ebkaj1 7	Clippe r (DBas e III)	0	10	Basic	0	5	C++	1	15
ebkaj1 8	Python	0	1	C#	0	0,5	C++	2	0
ebkaj2 0	C/C++	4	20	C55xx assembler	0	4			
ebkaj2 1	C++	0	2	C	2	2			

En general todos los individuos tienen experiencia en el lenguaje de programación C++ por lo tanto se pueden separar en dos grupos. Los novatos por un lado, y los intermedios y expertos por otro, considerando que los novatos están en desventaja frente a los intermedios y expertos. En la tabla 33 se muestran sombreados los individuos que se considera a priori que pueden influir negativamente en el experimento.

Tabla 103. Programación y experiencia en C++

ID	Programming experience	C++ experience
ebkaj11	Intermediate (5-<=10 years)	Intermediate (5-<=10 years)
ebkaj13	Novice (2-<=5 years)	Novice (2-<=5 years)
ebkaj14	Expert (>10 years)	Expert (>10 years)
ebkaj15	Intermediate (5-<=10 years)	Intermediate (5-<=10 years)
ebkaj16	Intermediate (5-<=10 years)	Novice (2-<=5 years)
ebkaj17	Expert (>10 years)	Expert (>10 years)
ebkaj18	Novice (2-<=5 years)	Novice (2-<=5 years)
ebkaj20	Expert (>10 years)	Expert (>10 years)
ebkaj21	Novice (2-<=5 years)	Novice (2-<=5 years)

4.3 EXPERIENCIA EN TESTING

La información de la tabla 34 presenta la experiencia de los individuos realizando pruebas de unidad. Más de la mitad de los individuos no tiene experiencia. Con estos datos podemos determinar que los individuos que no tienen ninguna experiencia se encuentran en desventaja con respecto al resto y por consiguiente influyen negativamente en el experimento.

Tabla 34. Experiencia en Unit testing

ID	Experiencia en Unit testing
ebkaj11	Intermediate (5-<=10 years)
ebkaj13	No experience (<2 years)
ebkaj14	Intermediate (5-<=10 years)
ebkaj15	Novice (2-<=5 years)
ebkaj16	No experience (<2 years)
ebkaj17	No experience (<2 years)
ebkaj18	No experience (<2 years)
ebkaj20	No experience (<2 years)
ebkaj21	Novice (2-<=5 years)

4.4 EXPERIENCIA EN FRAMEWORKS DE TESTING

Dado que la aplicación de TDD requiere una infraestructura de pruebas compuestas por determinadas herramientas, se considera que poseer experiencia previa en estas herramientas puede influir positivamente en el experimento. En concreto la experiencia en GTest es muy influyente en el experimento puesto que las pruebas son parte del ciclo de TDD.

En general todos los individuos no poseen experiencia en GTest por lo tanto se considera que esta variable no es influyente en los resultados del experimento.

Tabla 35. Unit testing tools and GTest experience

ID	Unit test tools used	GTest Experience
ebkaj11	cpputest, gcov	No experience (<2 years)
ebkaj13	N/A	No experience (<2 years)
ebkaj14	CppUTest	No experience (<2 years)
ebkaj15	some kind unit test tool	No experience (<2 years)
ebkaj16	utest	No experience (<2 years)
ebkaj17	no	No experience (<2 years)
ebkaj18	No	No experience (<2 years)
ebkaj20	no	No experience (<2 years)
ebkaj21	CPPUtest	No experience (<2 years)

4.5 USO DE IDES

El IDE que se usa en el experimento es Eclipse, se considera de mucha relevancia conocer el entorno de desarrollo puesto que puede facilitar mucho el trabajo de codificación ya que proporciona muchas utilidades que ahorran tiempo durante la programación.

La tabla 36 recoge la información relativa al uso de IDEs por parte de los participantes del experimento. En su mayoría están familiarizados con el entorno de desarrollo Eclipse, sin embargo hay tres individuos que no poseen experiencia usando Eclipse o cualquier otro IDE similar, por tanto se considera que influyen negativamente en el experimento.

Tabla 36. Uso de IDEs

ID	IDEs used
ebkaj11	eclipse cdt
ebkaj13	N/A
ebkaj14	c++/Eclipse
ebkaj15	-
ebkaj16	ecipse
ebkaj17	Eclipse, CodeBlocks
ebkaj18	-
ebkaj20	Eclipse
ebkaj21	Eclipse

4.6 EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO

La tabla 37 muestra la experiencia de los individuos en metodologías de desarrollo. Entendemos que la experiencia en metodologías ágiles es un aspecto positivo a tener en cuenta. Todos los individuos salvo ebkaj13 presentan experiencia en desarrollo con metodologías ágiles. Por tanto se considera que este individuo está en desventaja frente al resto de individuos y puede influir negativamente en el experimento.

Tabla 37. Metodologías de desarrollo

ID	Most recently used			Intermediate			Leastrecently used		
	Methd.	Educ. /yr	Usage /yr.	Methd.	Edu c. /yr	Usag e /yr.	Methd .	Edu c. /yr	Usag e /yr.
ebkaj21	Agile, Scrum	0	2						
ebkaj13	Waterfall	3 years of professional career	Iterative	1 years of professional career					
ebkaj16	iterative	1	5	waterfall	1	1	agile	0	2
ebkaj11	waterfall	0,5	3	agile, scrum	0,5	4	iterative	0,5	1
ebkaj18	agile, scrum	0	1						
ebkaj14	scrum	0 years	4 years	waterfall		9 years			
ebkaj15	waterfall	2	4	agile	2	4			
ebkaj17	Waterfall	3	10	Iterative	2	5	Agile Scrum		5
ebkaj20	agile (scrum)	0	2	waterfall	1	15			

4.7 EXPERIENCIA EN TDD

La experiencia en TDD es importante porque puede influir en el experimento. Si lo que queremos es comparar TDD con las metodologías tradicionales, el conocimiento que tengan sobre TDD de los sujetos que realizan el experimento podría ser una covariable.

La tabla 38 muestra que los encuestados en su mayoría no están familiarizados con esta técnica. Por lo tanto no se puede considerar una variable moderadora.

Table 38. Uso de TDD

ID	Usan TDD	Experiencia en TDD	Usan Katas	Experiencia usando Katas
ebkaj11	No		No	
ebkaj13	No		No	
ebkaj14	No		No	
ebkaj15	No	No experience (<2 years)	No	
ebkaj16	No		No	
ebkaj17	No		No	
ebkaj18	No		No	
ebkaj20	No		No	
ebkaj21	No		No	

El uso de Katas ayuda a los programadores a desarrollar y perfeccionar sus habilidades de programación a través de la práctica y la repetición. Se ha considerado que este factor puede influir a la hora de realizar el experimento. La tabla 38 muestra los resultados de los formularios. Como se puede ver ningún usuarios usa Katas con lo cual esta variable no será tomada en cuenta como variable influyente.

4.8 RESUMEN DE LAS POSIBLES VARIABLES MODERADORAS (CO-VARIABLES)

En la tabla 39 se recoge un resumen de la influencia de las variables que se han considerado moderadoras. Un signo “+” indica que pueden influir positivamente mientras que un signo “-” indica que puede influir negativamente en el experimento.

Hay otras variables que se han analizado y que sin embargo no aparecen en la tabla, esto es porque no se consideran influyentes en el experimento. Estas variables son: La experiencia usando frameworks de testing, el uso de TDD y el uso de Katas.

Tabla 39. Resumen de las posibles variables moderadoras

ID	Culificación y Experiencia	Experiencia en programación C++	Experiencia en pruebas de unidad	Uso de IDE	Metodologías de desarrollo
ebkaj11	+	+	+	+	+
ebkaj13	+	-	-	-	-
ebkaj14	+	+	+	+	+
ebkaj15	+	+	+	-	+
ebkaj16	+	-	-	+	+
ebkaj17	+	+	-	+	+
ebkaj18	-	-	-	-	+
ebkaj20	+	+	-	+	+
ebkaj21	-	-	+	+	+

4.9 CUANTIFICACIÓN DE LAS VARIABLES MODERADORAS

Con el objetivo de tener valores numéricos para poder usarlos como entrada de herramientas de análisis se elabora la tabla 40, contiene las variables moderadoras de la tabla 39 pero se ha proporcionado un valor numérico. A continuación se explica el criterio que se ha seguido para cuantificar cada variable:

4.9.1 CUALIFICACIÓN Y EXPERIENCIA

Esta variable se cuantifica teniendo en cuenta la preparación académica y la experiencia en la participación en proyectos software. Los valores de la tabla se corresponden con los años de experiencia de los individuos.

4.9.2 EXPERIENCIA DE PROGRAMACIÓN

Esta variable se cuantifica teniendo en cuenta sobre todo la experiencia que tienen los usuarios siendo desarrolladores. Los valores de la tabla se corresponden con los años de experiencia como desarrolladores usando el lenguaje de programación Java de cada uno de los usuarios. Se considera que lo importante es la experiencia que tienen los individuos en el desarrollo Java.

4.9.3 EXPERIENCIA EN PRUEBAS DE UNIDAD

Los valores de esta variable se corresponden con los años de experiencia de los usuarios realizando pruebas de unidad.

4.9.4 USO DE IDE

Para cuantificar esta variable se tiene en cuenta si están familiarizados con IDEs tipo Netbeans, Eclipse, VisualStudio, etc. Dado que no se puede cuantificar con un valor

natural se va a establecer un valor booleano. De esta forma la traducción es: los usuarios que influyen positivamente le corresponde el valor 1 y a lo que influyen negativamente les corresponde el valor 0.

4.9.5 METODOLOGÍAS DE DESARROLLO

Para cuantificar esta variable se usa la suma de los años de experiencia usando metodologías ágiles de desarrollo.

Table 40. Cuantificación de las variables moderadoras

ID	Culificación y Experiencia	Experiencia en programación C++	Experiencia en pruebas de unidad	Uso de IDE	Metodologías de desarrollo
ebkaj11	8	5	5	1	2
ebkaj13	2	2	0	0	0
ebkaj14	13	10	5	1	2
ebkaj15	7	5	2	0	4
ebkaj16	4	2	0	1	1
ebkaj17	15	10	0	1	4
ebkaj18	0	2	0	0	4
ebkaj20	20	10	0	1	5
ebkaj21	0	2	2	1	2

ANEXO B. Análisis Dicotómico

1 ANÁLISIS DICOTÓMICO MARS ROVER

El presente documento contiene un análisis que permite dilucidar la posible influencia de determinadas variables con respecto a las variables respuesta. Es decir, se intenta dar luz para poder determinar, en el caso de que sea posible, si se tratan de covariables del experimento.

Se realizaron tres ejecuciones del experimento, estas son:

1. MARS ROVER
2. BOWLING SCOREKEEPER
3. MUSICPHONE

Las variables respuesta son las siguientes:

- Productivity
- Quality
- TUS

Las posibles covariables que se analizan son:

- Experiencia académica
- Experiencia de programación
- Experiencia de Java
- Experiencia en Testing
- Experiencia en el uso de JUnit
- Experiencia en el uso de IDEs
- Experiencia en metodologías de desarrollo
- Experiencia en el uso TDD

A continuación se analizan las posibles covariables frente a cada una de las variables respuesta para cada uno de las ejecuciones de los experimentos.

1.1 PRODUCTIVIDAD - EXPERIENCIA ACADEMICA

Para analizar la correlación entre la experiencia profesional y la productividad lo primero que hacemos es el test de normalidad para saber que análisis estadístico se puede aplicar.

1.1.1 PRUEBA DE NORMALIDAD

Dado que se trata de una muestra de menos de 30 individuos, nos fijamos en los resultados de Shapiro-Wilk. Como se puede ver en la tabla 1, la variable experiencia académica no sigue una distribución normal.

Pruebas de normalidad							
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,441	4	.	,630	4	,001
	Positiva	,187	20	,066	,896	20	,035

a. Corrección de la significación de Lilliefors

Tabla 1: Prueba de normalidad para la variable "experiencia académica"

Para los datos negativos la significancia es de $0,001 < 0,05$ y para los valores positivos es $0,035 < 0,05$ por lo tanto los datos de la variable "Experiencia académica" no superan el test de Normalidad dado que existe significancia como para rechazar la hipótesis nula de que hay normalidad.

1.1.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Academic Exp				Estadístico	Error típ.
Productivity	Negativa	Media		8,9075	8,90750
		Intervalo de confianza para la media al 95%	Límite inferior	-19,4401	
			Límite superior	37,2551	
		Media recortada al 5%		7,9178	
		Mediana		,0000	
		Varianza		317,374	
		Desv. típ.		17,81500	
		Mínimo		,00	
		Máximo		35,63	
		Rango		35,63	
		Amplitud intercuartil		26,72	
		Asimetría		2,000	1,014
		Curtosis		4,000	2,619
	Positiva	Media		21,4360	4,15214
		Intervalo de confianza para la media al 95%	Límite inferior	12,7455	
			Límite superior	30,1265	
		Media recortada al 5%		20,2417	
		Mediana		14,3650	
		Varianza		344,806	
		Desv. típ.		18,56895	
		Mínimo		,00	
		Máximo		64,37	
		Rango		64,37	
		Amplitud intercuartil		33,33	
		Asimetría		,612	,512
		Curtosis		-,472	,992

Tabla 2: Estadísticos descriptivos para la variable "experiencia academica"

1.1.3 DIAGRAMA DE CAJAS: PRODUCTIVIDAD – EXPERIENCIA ACADÉMICA

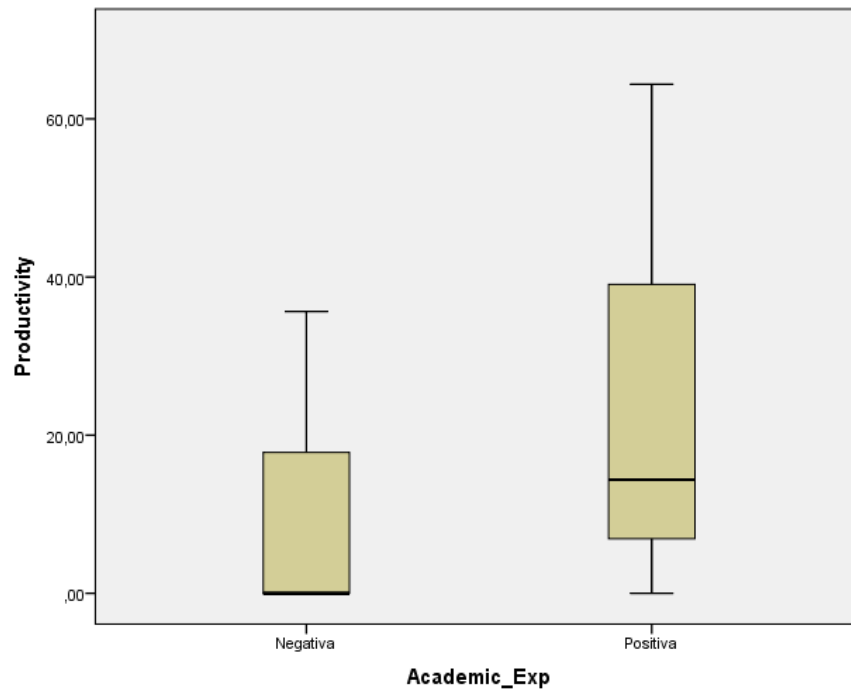


Tabla 3: Diagrama de cajas de la variable "Experiencia académica"

1.1.4 RESULTADOS

Dado que la variable “Experiencia académica” no supera el test de normalidad aplicamos una prueba no paramétrica, para comprobar que la distribución de la productividad es la misma entre las categorías de la variable “Experiencia académica” (Negativa y Positiva).

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Experience.	Prueba U de Mann-Whitney de muestras independientes	,135 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Tabla 4: Prueba de hipótesis para la variable "Experiencia académica"

Los resultados muestran que se acepta la hipótesis nula que sostiene que las categorías de la variable “Experiencia académica” siguen la misma distribución de productividad. Sin embargo el valor de la significancia es muy bajo, lo que se puede

explicar con que hay pocos datos para la categoría “Negativos”, solamente 4 como refleja la Tabla 1.

1.2 PRODUCTIVIDAD – EXPERIENCIA JAVA

1.2.1 PRUEBAS DE NORMALIDAD

Pruebas de normalidad

	Progr_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,441	4	.	,630	4	,001
	Positiva	,183	20	,078	,900	20	,042

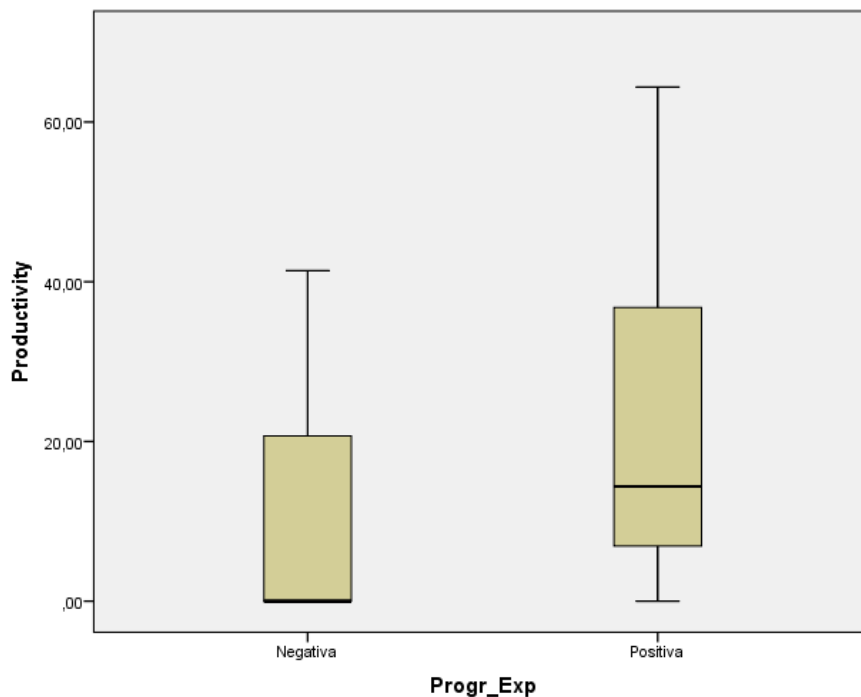
a. Corrección de la significación de Lilliefors

1.2.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Progr_Exp			Estadístico	Error típ.
Productivity	Negativa	Media	10,3450	10,34500
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	-22,5774 43,2674
		Media recortada al 5%	9,1956	
		Mediana	,0000	
		Varianza	428,076	
		Desv. típ.	20,69000	
		Mínimo	,00	
		Máximo	41,38	
		Rango	41,38	
		Amplitud intercuartil	31,04	
		Asimetría	2,000	1,014
		Curtosis	4,000	2,619
	Positiva	Media	21,1485	4,08893
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	12,5903 29,7067
		Media recortada al 5%	19,9222	
		Mediana	14,3650	
		Varianza	334,388	
		Desv. típ.	18,28627	
		Mínimo	,00	
		Máximo	64,37	
		Rango	64,37	
		Amplitud intercuartil	31,03	
		Asimetría	,645	,512
		Curtosis	-,306	,992

1.2.3 DIAGRAMA DE CAJAS: PRODUCTIVIDAD – EXPERIENCIA DE PROGRAMACIÓN



1.2.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Progr_Exp.	Prueba U de Mann-Whitney de muestras independientes	,241 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Los resultados muestran que se acepta la hipótesis nula que sostiene que las categorías de la variable “Experiencia de programación” siguen la misma distribución de productividad.

1.3 PRODUCTIVIDAD – EXPERIENCIA EN TDD

1.3.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	TDD_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,293	12	,005	,763	12	,004
	Positiva	,215	12	,131	,899	12	,153

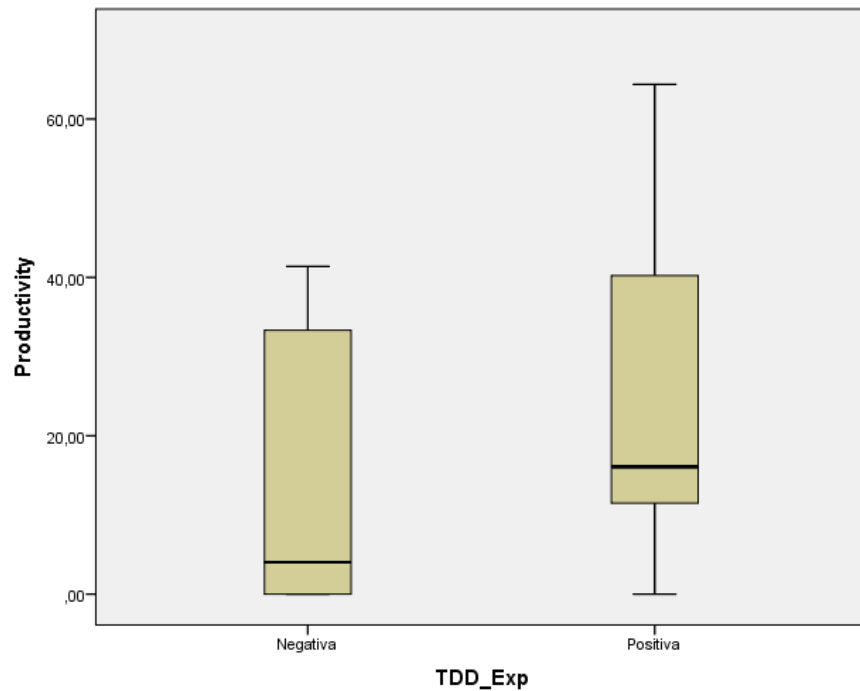
a. Corrección de la significación de Lilliefors

La variable experiencia en TDD no supera el test de normalidad.

1.3.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					Estadístico	Error típ.
TDD Exp						
Productivity	Negativa	Media			13,8883	4,90590
		Intervalo de confianza para la media al 95%	Límite inferior		3,0905	
			Límite superior		24,6861	
		Media recortada al 5%			13,1326	
		Mediana			4,0250	
		Varianza			288,814	
		Desv. típ.			16,99453	
		Mínimo			,00	
		Máximo			41,38	
		Rango			41,38	
		Amplitud intercuartil			33,33	
		Asimetría			,655	,637
		Curtosis			-1,568	1,232
	Positiva	Media			24,8075	5,59648
		Intervalo de confianza para la media al 95%	Límite inferior		12,4897	
			Límite superior		37,1253	
		Media recortada al 5%			23,9878	
		Mediana			16,0900	
		Varianza			375,847	
		Desv. típ.			19,38678	
		Mínimo			,00	
		Máximo			64,37	
		Rango			64,37	
		Amplitud intercuartil			29,31	
		Asimetría			,683	,637
		Curtosis			-,371	1,232

1.3.3 DIAGRAMA DE CAJAS: PRODUCTIVIDAD – EXPERIENCIA EN TDD



1.3.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de TDD_Exp.	Prueba U de Mann-Whitney de muestras independientes	,089 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.4 PRODUCTIVIDAD – EXPERIENCIA EN JUNIT

1.4.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Junit_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,355	9	,002	,687	9	,001
	Positiva	,220	15	,050	,887	15	,060

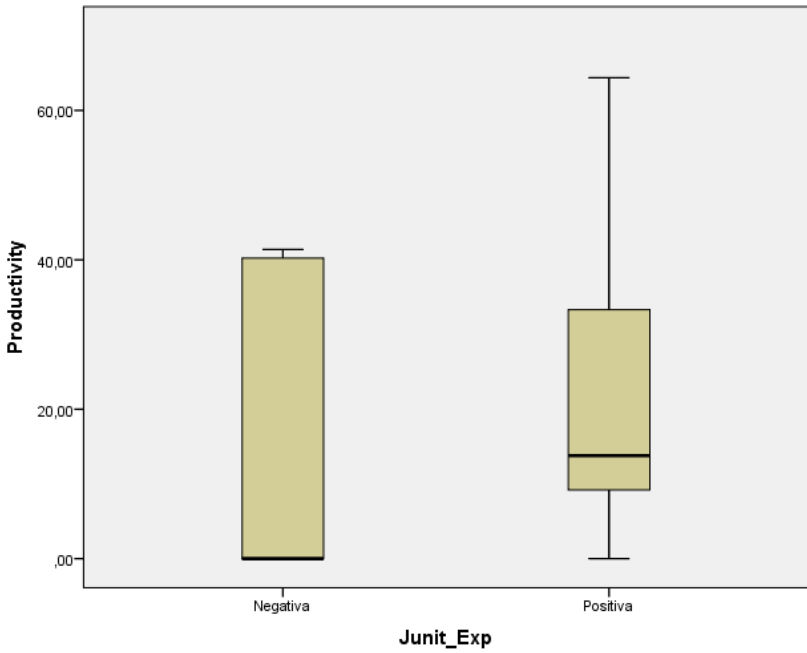
a. Corrección de la significación de Lilliefors

La variable experiencia en JUnit no supera el test de normalidad.

1.4.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Junit_Exp				Estadístico	Error típ.
Productivity	Negativa	Media		17,6244	6,98905
		Intervalo de confianza para la media al 95%	Límite inferior	1,5077	
			Límite superior	33,7412	
		Media recortada al 5%		17,2838	
		Mediana		,0000	
		Varianza		439,621	
		Desv. típ.		20,96714	
		Mínimo		,00	
		Máximo		41,38	
		Rango		41,38	
		Amplitud intercuartil		40,81	
		Asimetría		,293	,717
		Curtosis		-2,512	1,400
	Positiva	Media		20,3820	4,61205
		Intervalo de confianza para la media al 95%	Límite inferior	10,4901	
			Límite superior	30,2739	
		Media recortada al 5%		19,0706	
		Mediana		13,7900	
		Varianza		319,065	
		Desv. típ.		17,86238	
		Mínimo		,00	
		Máximo		64,37	
		Rango		64,37	
		Amplitud intercuartil		25,28	
		Asimetría		1,133	,580
		Curtosis		1,072	1,121

1.4.3 DIAGRAMA DE CAJAS: PRODUCTIVIDAD – EXPERIENCIA EN JUNIT



1.4.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Junit_Exp.	Prueba U de Mann-Whitney de muestras independientes	,599 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.5 PRODUCTIVIDAD – EXPERIENCIA IDE

1.5.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	IDE_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,375	4	.	,796	4	,095
	Positiva	,203	20	,031	,831	20	,003

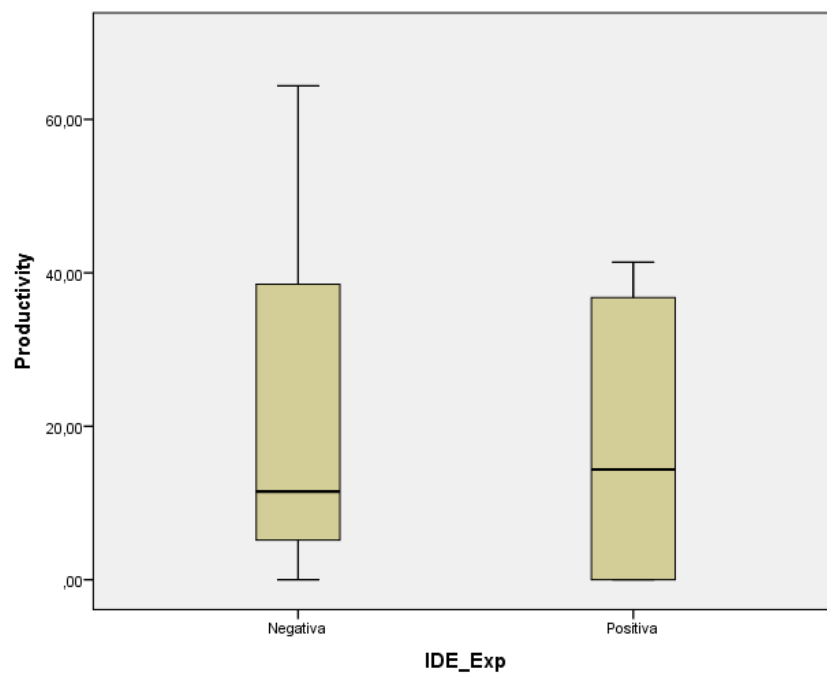
a. Corrección de la significación de Lilliefors

La variable Experiencia en el uso de IDEs no supera el test de normalidad.

1.5.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
IDE_Exp				Estadístico	Error típ.
Productivity	Negativa	Media		21,8375	14,44148
		Intervalo de confianza para la media al 95%	Límite inferior	-24,1217	
			Límite superior	67,7967	
		Media recortada al 5%		20,6878	
		Mediana		11,4900	
		Varianza		834,225	
		Desv. típ.		28,88296	
		Mínimo		,00	
		Máximo		64,37	
		Rango		64,37	
		Amplitud intercuartil		48,85	
		Asimetría		1,777	1,014
		Curtosis		3,382	2,619
	Positiva	Media		18,8500	3,80243
		Intervalo de confianza para la media al 95%	Límite inferior	10,8914	
			Límite superior	26,8086	
		Media recortada al 5%		18,6456	
		Mediana		14,3650	
		Varianza		289,169	
		Desv. típ.		17,00498	
		Mínimo		,00	
		Máximo		41,38	
		Rango		41,38	
		Amplitud intercuartil		37,36	
		Asimetría		,191	,512
		Curtosis		-1,780	,992

1.5.3 DIAGRAMA DE CAJAS: PRODUCTIVIDAD – EXPERIENCIA IDE



1.5.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de IDE_Exp.	Prueba U de Mann-Whitney de muestras independientes	1,000 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.6 PRODUCTIVIDAD – EXPERIENCIA PRUEBAS DE UNIDAD

1.6.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Unit_Test_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,338	8	,008	,663	8	,001
	Positiva	,164	16	,200 [*]	,925	16	,200

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

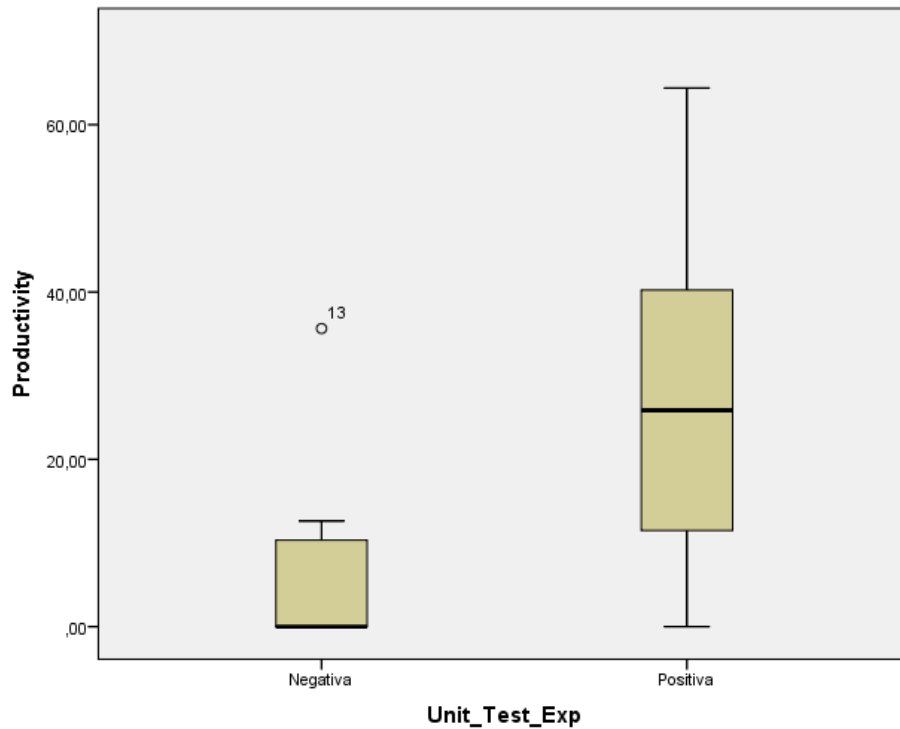
La variable experiencia en pruebas de unidad no supera el test de normalidad.

1.6.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Unit Test Exp				Estadístico	Error típ.
Productivity	Negativa	Media		7,0400	4,42714
		Intervalo de confianza para la media al 95%	Límite inferior	-3,4285	
			Límite superior	17,5085	
		Media recortada al 5%		5,8428	
		Mediana		,0000	
		Varianza		156,797	
		Desv. típ.		12,52184	
		Mínimo		,00	
		Máximo		35,63	
		Rango		35,63	
		Amplitud intercuartil		11,49	
		Asimetría		2,115	,752
		Curtosis		4,602	1,481
	Positiva	Media		25,5019	4,61485
		Intervalo de confianza para la media al 95%	Límite inferior	15,6655	
			Límite superior	35,3382	
		Media recortada al 5%		24,7593	
		Mediana		25,8600	
		Varianza		340,750	
		Desv. típ.		18,45942	
		Mínimo		,00	
		Máximo		64,37	
		Rango		64,37	
		Amplitud intercuartil		29,31	
		Asimetría		,319	,564
		Curtosis		-,579	1,091

1.6.3 DIAGRAMA DE CAJAS: PRODUCTIVIDAD



1.6.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Unit_Test_Exp.	Prueba U de Mann-Whitney de muestras independientes	,009 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.7 PRODUCTIVIDAD – EXPERIENCIA DE PROGRAMACIÓN

1.7.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Programacion_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,356	4	.	,740	4	,031
	Positiva	,179	20	,093	,890	20	,027

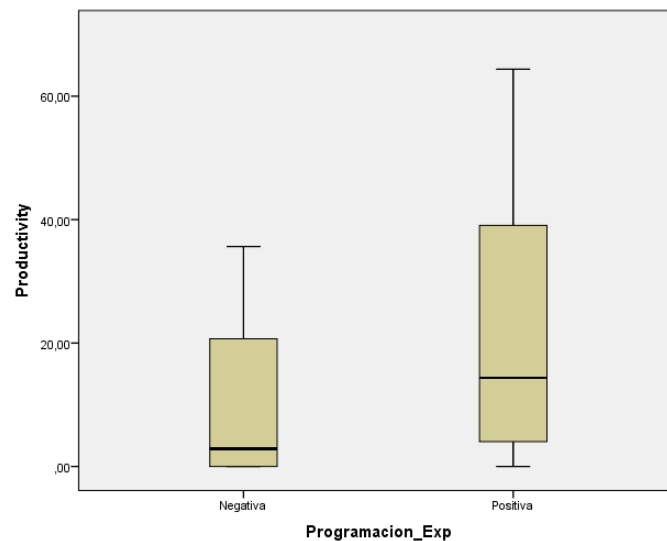
a. Corrección de la significación de Lilliefors

La variable Experiencia de programación no supera el test de normalidad.

1.7.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Programacion Exp				Estadístico	Error típ.
Productivity	Negativa	Media		10,3450	8,53660
		Intervalo de confianza para la media al 95%	Límite inferior	-16,8223	
			Límite superior	37,5123	
		Media recortada al 5%		9,5150	
		Mediana		2,8750	
		Varianza		291,494	
		Desv. típ.		17,07321	
		Mínimo		,00	
		Máximo		35,63	
		Rango		35,63	
		Amplitud intercuartil		28,16	
		Asimetría		1,856	1,014
		Curtosis		3,451	2,619
	Positiva	Media		21,1485	4,21873
		Intervalo de confianza para la media al 95%	Límite inferior	12,3186	
			Límite superior	29,9784	
		Media recortada al 5%		19,9222	
		Mediana		14,3650	
		Varianza		355,953	
		Desv. típ.		18,86672	
		Mínimo		,00	
		Máximo		64,37	
		Rango		64,37	
		Amplitud intercuartil		37,64	
		Asimetría		,582	,512
		Curtosis		-,534	,992

1.7.3 DIAGRAMA DE CAJAS



1.7.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Programacion_Exp.	Prueba U de Mann-Whitney de muestras independientes	,241 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.8 PRODUCTIVIDAD – METODOLOGÍAS DE DESARROLLO

1.8.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	MetDes_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,243	5	,200*	,889	5	,354
	Positiva	,185	19	,086	,853	19	,008

*. Este es un límite inferior de la significación verdadera.

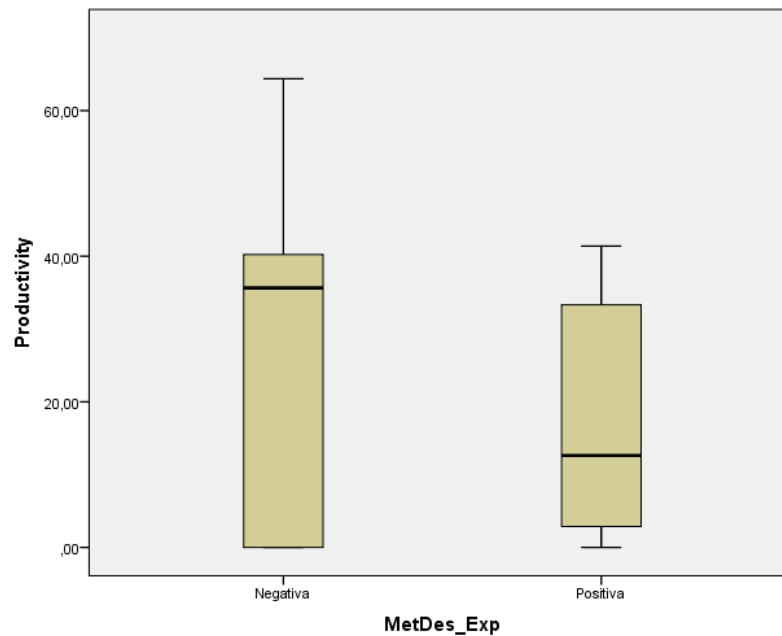
a. Corrección de la significación de Lilliefors

La variable Metodologías de desarrollo no supera el test de normalidad.

1.8.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				Estadístico	Error típ.
Productivity	MetDes	Exp			
Productivity	Negativa	Media		28,0460	12,44700
		Intervalo de confianza para la media al 95%	Límite inferior	-6,5124	
			Límite superior	62,6044	
		Media recortada al 5%		27,5861	
		Mediana		35,6300	
		Varianza		774,639	
		Desv. típ.		27,83234	
		Mínimo		,00	
		Máximo		64,37	
		Rango		64,37	
		Amplitud intercuartil		52,30	
		Asimetría		,117	,913
		Curtosis		-1,743	2,000
	Positiva	Media		17,0589	3,60728
		Intervalo de confianza para la media al 95%	Límite inferior	9,4803	
			Límite superior	24,6376	
		Media recortada al 5%		16,6555	
		Mediana		12,6400	
		Varianza		247,237	
		Desv. típ.		15,72376	
		Mínimo		,00	
		Máximo		41,38	
		Rango		41,38	
		Amplitud intercuartil		33,33	
		Asimetría		,502	,524
		Curtosis		-1,330	1,014

1.8.3 DIAGRAMA DE CAJAS



1.8.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de MetDes_Exp.	Prueba U de Mann-Whitney de muestras independientes	,581 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.9 CALIDAD – EXPERIENCIA ACADÉMICA

1.9.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Academic_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,441	4	.	,630	4	,001
	Positiva	,194	20	,048	,811	20	,001

a. Corrección de la significación de Lilliefors

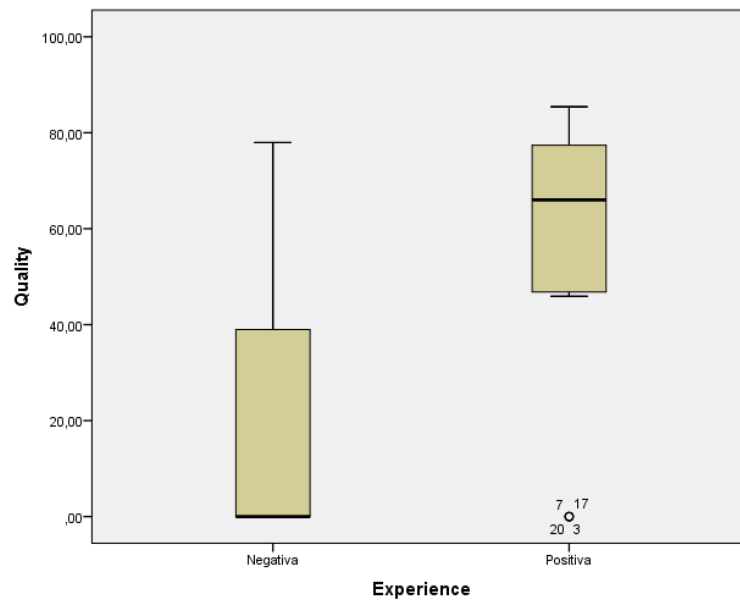
La variable experiencia en académica no supera el test de normalidad.

1.9.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Academic Exp			Estadístico	Error típ.
Quality	Negativa	Media	19,4875	19,48750
		Intervalo de confianza para la media al 95%	Límite inferior	-42,5304
			Límite superior	81,5054
		Media recortada al 5%	17,3222	
		Mediana	,0000	
		Varianza	1519,051	
		Desv. típ.	38,97500	
		Mínimo	,00	
		Máximo	77,95	
		Rango	77,95	
		Amplitud intercuartil	58,46	
		Asimetría	2,000	1,014
		Curtosis	4,000	2,619
	Positiva	Media	54,0735	6,76322
		Intervalo de confianza para la media al 95%	Límite inferior	39,9179
			Límite superior	68,2291
		Media recortada al 5%	55,3361	
		Mediana	66,0150	
		Varianza	914,822	
		Desv. típ.	30,24602	
		Mínimo	,00	
		Máximo	85,42	
		Rango	85,42	
		Amplitud intercuartil	31,08	
		Asimetría	-1,038	,512
		Curtosis	-,254	,992

1.9.3 DIAGRAMA DE CAJAS: PRODUCTIVIDAD



1.9.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de Academic_Exp.	Prueba U de Mann-Whitney de muestras independientes	,183 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.10 CALIDAD – EXPERIENCIA JAVA

1.10.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Progr_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,441	4	.	,630	4	,001
	Positiva	,197	20	,041	,806	20	,001

a. Corrección de la significación de Lilliefors

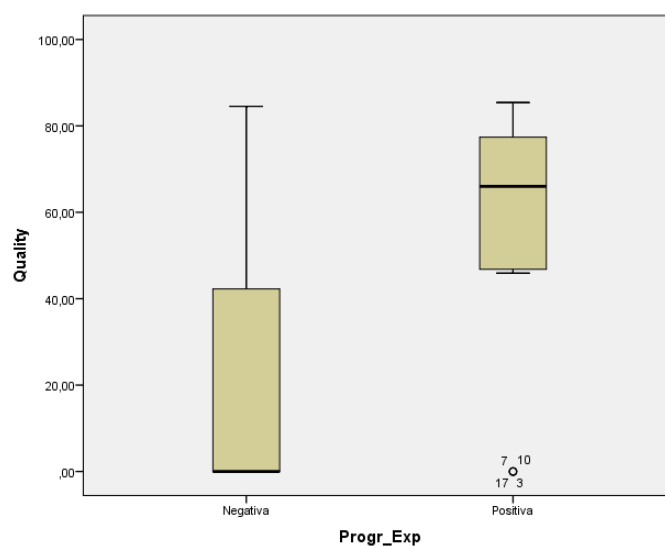
La variable experiencia de programación no supera el test de normalidad.

1.10.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Progr. Exp			Estadístico	Error típ.
Quality	Negativa	Media	21,1300	21,13000
		Intervalo de confianza para la media al 95%	Límite inferior	-46,1151
			Límite superior	88,3751
		Media recortada al 5%	18,7822	
		Mediana	,0000	
		Varianza	1785,908	
		Desv. típ.	42,26000	
		Mínimo	,00	
		Máximo	84,52	
		Rango	84,52	
		Amplitud intercuartil	63,39	
		Asimetría	2,000	1,014
		Curtosis	4,000	2,619
	Positiva	Media	53,7450	6,69300
		Intervalo de confianza para la media al 95%	Límite inferior	39,7364
			Límite superior	67,7536
		Media recortada al 5%	54,9711	
		Mediana	66,0150	
		Varianza	895,924	
		Desv. típ.	29,93199	
		Mínimo	,00	
		Máximo	85,42	
		Rango	85,42	
		Amplitud intercuartil	31,08	
		Asimetría	-1,066	,512
		Curtosis	-,221	,992

1.10.3 DIAGRAMA DE CAJAS: PRODUCTIVIDAD



1.10.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de Progr_Exp.	Prueba U de Mann-Whitney de muestras independientes	,273 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.11 CALIDAD – EXPERIENCIA EN TDD

1.11.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	TDD_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,324	12	,001	,748	12	,003
	Positiva	,198	12	,200*	,840	12	,028

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

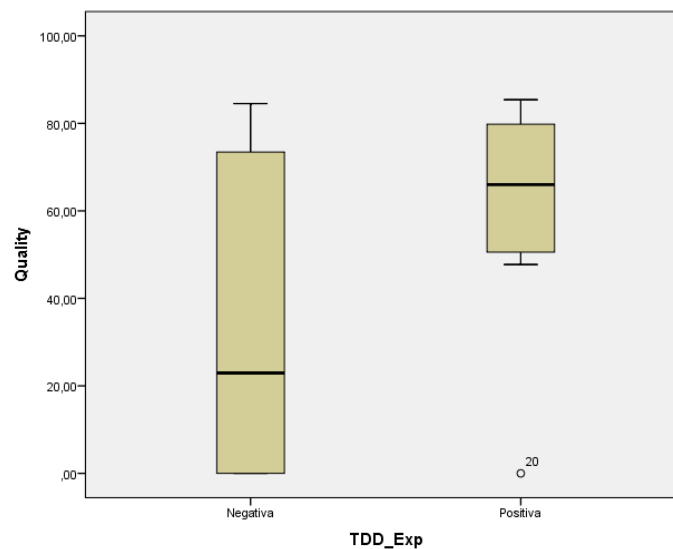
La variable experiencia en TDD no supera el test de normalidad.

1.11.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

TDD_Exp			Estadístico	Error típ.
Quality	Negativa	Media	35,4108	10,99261
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	11,2163 59,6054
		Media recortada al 5%	34,6498	
		Mediana	22,9550	
		Varianza	1450,050	
		Desv. típ.	38,07952	
		Mínimo	,00	
		Máximo	84,52	
		Rango	84,52	
		Amplitud intercuartil	75,36	
		Asimetría	,165	,637
		Curtosis	-2,208	1,232
	Positiva	Media	61,2075	6,73228
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	46,3898 76,0252
		Media recortada al 5%	63,2628	
		Mediana	66,0150	
		Varianza	543,883	
		Desv. típ.	23,32131	
		Mínimo	,00	
		Máximo	85,42	
		Rango	85,42	
		Amplitud intercuartil	30,98	
		Asimetría	-1,693	,637
		Curtosis	3,902	1,232

1.11.3 DIAGRAMA DE CAJAS: PRODUCTIVIDAD



1.11.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de TDD_Exp.	Prueba U de Mann-Whitney de muestras independientes	,178 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.12 CALIDAD – EXPERIENCIA EN JUNIT

1.12.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Junit_Exp	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,356	9	,002	,678	9	,001
	Positiva	,221	15	,047	,808	15	,005

a. Corrección de la significación de Lilliefors

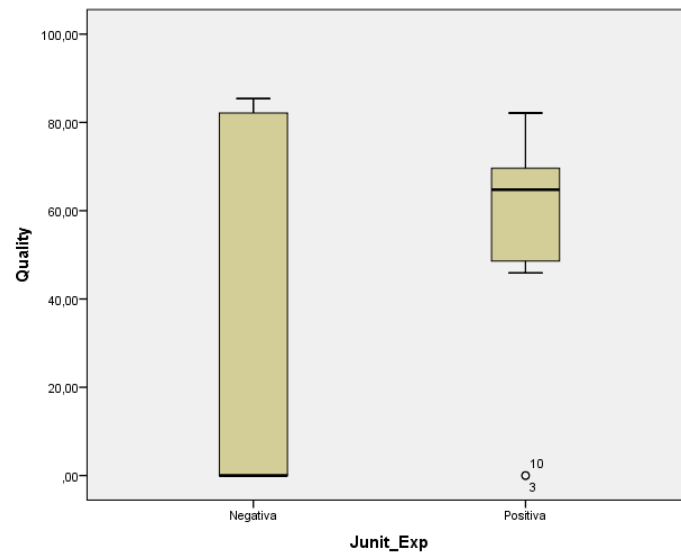
La variable experiencia en TDD no supera el test de normalidad.

1.12.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Junit_Exp			Estadístico	Error típ.
Quality	Negativa	Media	36,6700	14,51110
		Intervalo de confianza para la media al 95%	Límite inferior 3,2073 Límite superior 70,1327	
		Media recortada al 5%	35,9989	
		Mediana	,0000	
		Varianza	1895,149	
		Desv. típ.	43,53330	
		Mínimo	,00	
		Máximo	85,42	
		Rango	85,42	
		Amplitud intercuartil	83,33	
		Asimetría	,279	,717
		Curtosis	-2,550	1,400
	Positiva	Media	55,2927	6,47858
		Intervalo de confianza para la media al 95%	Límite inferior 41,3975 Límite superior 69,1878	
		Media recortada al 5%	56,8730	
		Mediana	64,7700	
		Varianza	629,580	
		Desv. típ.	25,09142	
		Mínimo	,00	
		Máximo	82,14	
		Rango	82,14	
		Amplitud intercuartil	21,91	
		Asimetría	-1,504	,580
		Curtosis	1,731	1,121

1.12.3 DIAGRAMA DE CAJAS: CALIDAD



1.12.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de JUnit_Exp.	Prueba U de Mann-Whitney de muestras independientes	,815 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Los sujetos que poseen experiencia en el uso de JUnit presentan una productividad similar, sin embargo los que no están familiarizados con JUnit tienen un rango mayor de productividad, están más dispersos, presentando en su mayoría menos Calidad que los que si poseen experiencia en JUnit.

1.13 CALIDAD – EXPERIENCIA IDE

1.13.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	IDE_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,385	4	.	,727	4	,023
	Positiva	,219	20	,013	,798	20	,001

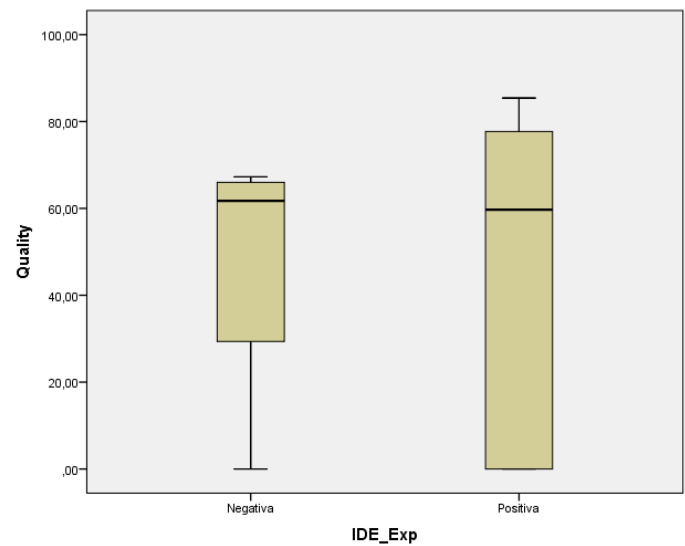
a. Corrección de la significación de Lilliefors

La variable experiencia en el uso de IDEs no supera el test de normalidad.

1.13.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
IDE_Exp				Estadístico	Error típ.
Quality	Negativa	Media		47,6850	15,99606
		Intervalo de confianza para la media al 95%	Límite inferior	-3,2216	
			Límite superior	98,5916	
		Media recortada al 5%		49,2467	
		Mediana		61,7400	
		Varianza		1023,496	
		Desv. típ.		31,99212	
		Mínimo		,00	
		Máximo		67,26	
		Rango		67,26	
		Amplitud intercuartil		51,96	
		Asimetría		-1,926	1,014
		Curtosis		3,738	2,619
	Positiva	Media		48,4340	7,75509
		Intervalo de confianza para la media al 95%	Límite inferior	32,2024	
			Límite superior	64,6656	
		Media recortada al 5%		49,0700	
		Mediana		59,6950	
		Varianza		1202,827	
		Desv. típ.		34,68180	
		Mínimo		,00	
		Máximo		85,42	
		Rango		85,42	
		Amplitud intercuartil		77,84	
		Asimetría		-,571	,512
		Curtosis		-1,448	,992

1.13.3 DIAGRAMA DE CAJAS: CALIDAD



1.13.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de IDE_Exp.	Prueba U de Mann-Whitney de muestras independientes	,627 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.14 CALIDAD – EXPERIENCIA PRUEBAS DE UNIDAD

1.14.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Unit_Test_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,384	8	,001	,696	8	,002
	Positiva	,201	16	,082	,807	16	,003

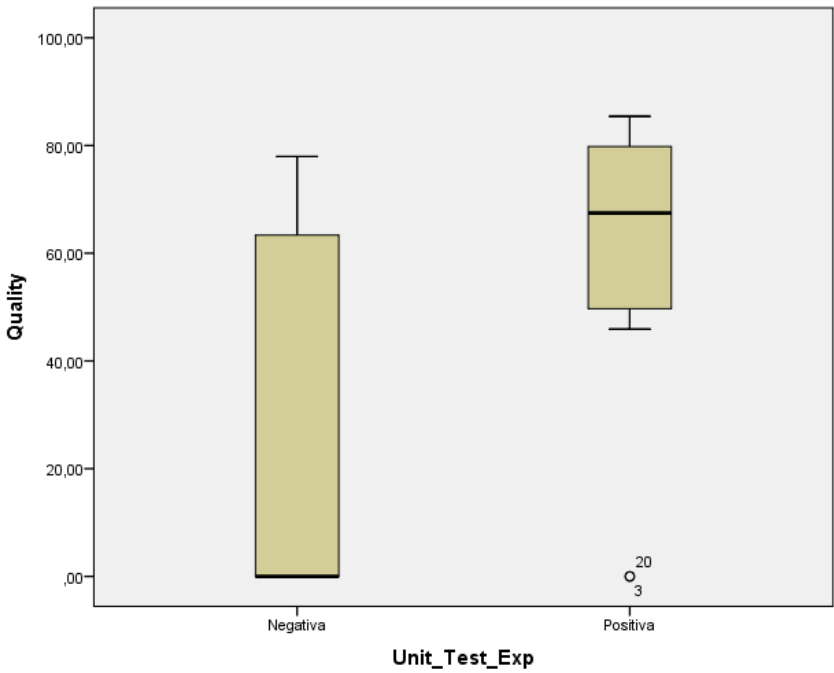
a. Corrección de la significación de Lilliefors

La variable experiencia en pruebas de unidad no supera el test de normalidad.

1.14.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Unit Test Exp			Estadístico	Error típ.
Quality	Negativa	Media	25,5812	12,85564
		Intervalo de confianza para la media al 95%	Límite inferior	-4,8175
			Límite superior	55,9800
		Media recortada al 5%	24,0931	
		Mediana	,0000	
		Varianza	1322,140	
		Desv. típ.	36,36124	
		Mínimo	,00	
		Máximo	77,95	
		Rango	77,95	
		Amplitud intercuartil	70,31	
		Asimetría	,838	,752
		Curtosis	-1,541	1,481
	Positiva	Media	59,6731	6,60568
		Intervalo de confianza para la media al 95%	Límite inferior	45,5935
			Límite superior	73,7528
		Media recortada al 5%	61,5579	
		Mediana	67,4750	
		Varianza	698,160	
		Desv. típ.	26,42271	
		Mínimo	,00	
		Máximo	85,42	
		Rango	85,42	
		Amplitud intercuartil	32,26	
		Asimetría	-1,512	,564
		Curtosis	1,785	1,091

1.14.3 DIAGRAMA DE CAJAS: CALIDAD



1.14.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de Unit_Test_Exp.	Prueba U de Mann-Whitney de muestras independientes	,052 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La experiencia en pruebas de unidad presenta una influencia clara en la Calidad.

1.15 CALIDAD – EXPERIENCIA DE PROGRAMACIÓN

1.15.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Programacion_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,294	4	.	,859	4	,256
	Positiva	,206	20	,026	,797	20	,001

a. Corrección de la significación de Lilliefors

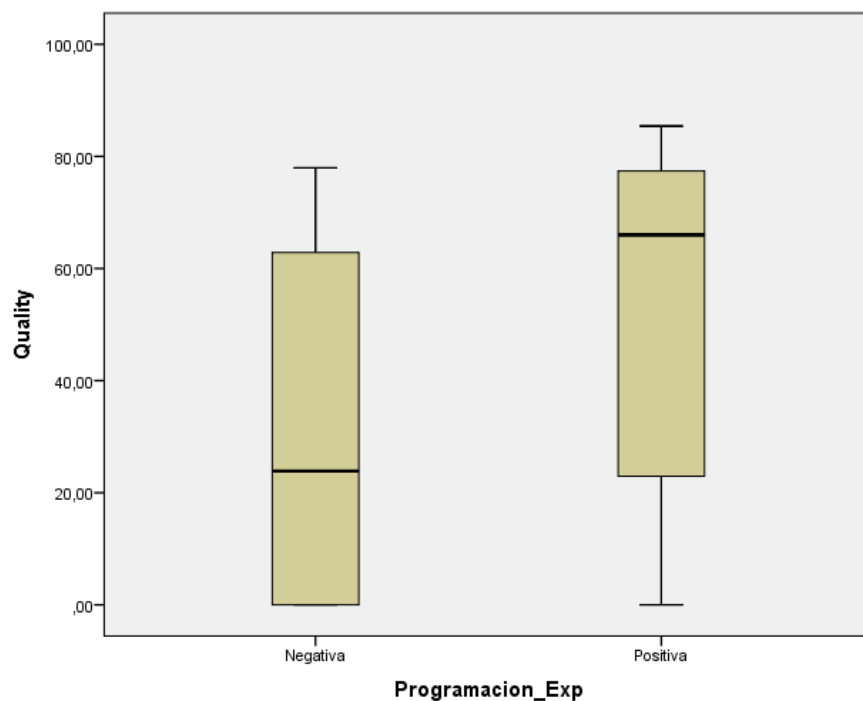
La variable Experiencia de programación no supera el test de normalidad.

1.15.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Programacion_Exp			Estadístico	Error típ.
Quality	Negativa	Media	31,4200	19,16048
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	-29,5572 92,3972
		Media recortada al 5%	30,5806	
		Mediana	23,8650	
		Varianza	1468,497	
		Desv. típ.	38,32097	
		Mínimo	,00	
		Máximo	77,95	
		Rango	77,95	
		Amplitud intercuartil	70,40	
		Asimetría	,510	1,014
		Curtosis	-3,132	2,619
	Positiva	Media	51,6870	7,28217
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	36,4452 66,9288
		Media recortada al 5%	52,6844	
		Mediana	66,0150	
		Varianza	1060,601	
		Desv. típ.	32,56686	
		Mínimo	,00	
		Máximo	85,42	
		Rango	85,42	
		Amplitud intercuartil	65,97	
		Asimetría	-,854	,512
		Curtosis	-,900	,992

1.15.3 DIAGRAMA DE CAJAS



1.15.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de Programacion_Exp.	Prueba U de Mann-Whitney de muestras independientes	,347 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹ Se muestra la significancia exacta para esta prueba.

1.16 CALIDAD – METODOLOGÍAS DE DESARROLLO

1.16.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	MetDes_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,299	5	,166	,769	5	,044
	Positiva	,198	19	,048	,823	19	,003

a. Corrección de la significación de Lilliefors

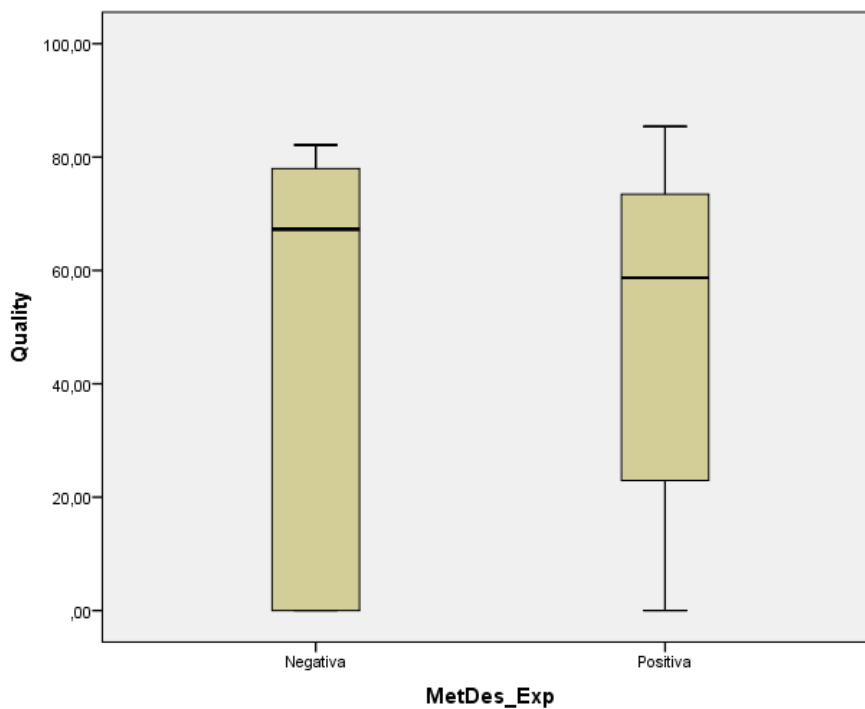
La variable Metodologías de desarrollo no supera el test de normalidad.

1.16.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

MetDes Exp			Estadístico	Error típ.
Quality	Negativa	Media	45,4700	18,72096
		Intervalo de confianza para la media al 95%	Límite inferior	-6,5077
			Límite superior	97,4477
		Media recortada al 5%	45,9589	
		Mediana	67,2600	
		Varianza	1752,371	
		Desv. típ.	41,86133	
		Mínimo	,00	
		Máximo	82,14	
		Rango	82,14	
		Amplitud intercuartil	80,05	
		Asimetría	-,534	,913
		Curtosis	-3,239	2,000
	Positiva	Media	49,0563	7,42717
		Intervalo de confianza para la media al 95%	Límite inferior	33,4524
			Límite superior	64,6602
		Media recortada al 5%	49,7615	
		Mediana	58,7100	
		Varianza	1048,093	
		Desv. típ.	32,37427	
		Mínimo	,00	
		Máximo	85,42	
		Rango	85,42	
		Amplitud intercuartil	77,27	
		Asimetría	-,709	,524
		Curtosis	-1,078	1,014

1.16.3 DIAGRAMA DE CAJAS



1.16.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de MetDes_Exp.	Prueba U de Mann-Whitney de muestras independientes	,945 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.17 TUS – EXPERIENCIA ACADÉMICA

1.17.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Academic_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,441	4	.	,630	4	,001
	Positiva	,185	20	,070	,904	20	,048

a. Corrección de la significación de Lilliefors

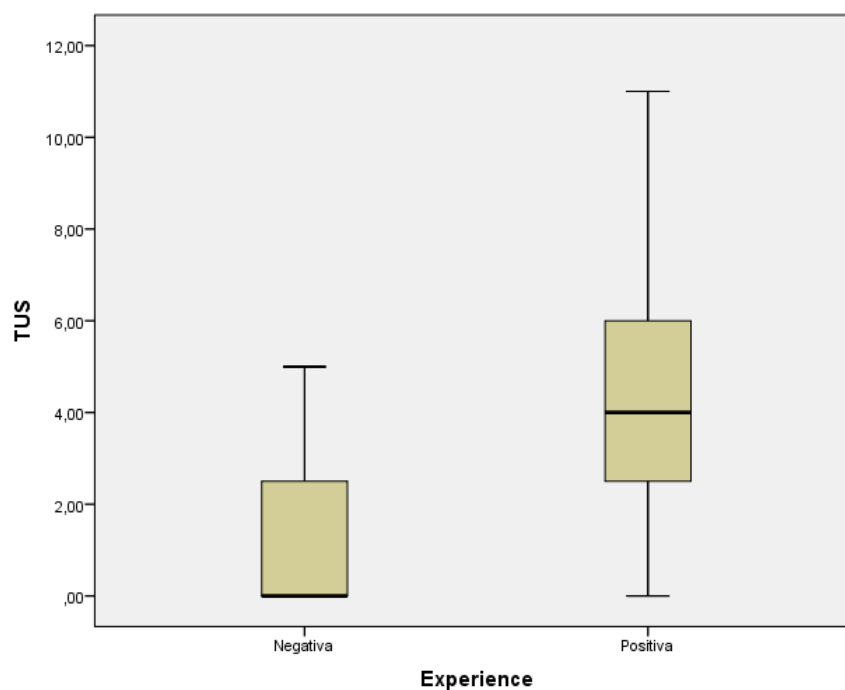
La variable experiencia en académica no supera el test de normalidad.

1.17.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Academic Exp				Estadístico	Error tip.
TUS	Negativa	Media		1,2500	1,25000
		Intervalo de confianza para la media al 95%	Límite inferior	-2,7281	
			Límite superior	5,2281	
		Media recortada al 5%		1,1111	
		Mediana		,0000	
		Varianza		6,250	
		Desv. típ.		2,50000	
		Mínimo		,00	
		Máximo		5,00	
		Rango		5,00	
		Amplitud intercuartil		3,75	
		Asimetría		2,000	1,014
		Curtosis		4,000	2,619
	Positiva	Media		4,0000	,61985
		Intervalo de confianza para la media al 95%	Límite inferior	2,7026	
			Límite superior	5,2974	
		Media recortada al 5%		3,8333	
		Mediana		4,0000	
		Varianza		7,684	
		Desv. típ.		2,77204	
		Mínimo		,00	
		Máximo		11,00	
		Rango		11,00	
		Amplitud intercuartil		3,75	
		Asimetría		,346	,512
		Curtosis		,794	,992

1.17.3 DIAGRAMA DE CAJAS: TUS



1.17.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Academic_Exp.	Prueba U de Mann-Whitney de muestras independientes	,097 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La experiencia académica influye claramente en la variable respuesta: TUS.

1.18 TUS – EXPERIENCIA JAVA

1.18.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Progr_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,441	4	.	,630	4	,001
	Positiva	,177	20	,099	,907	20	,055

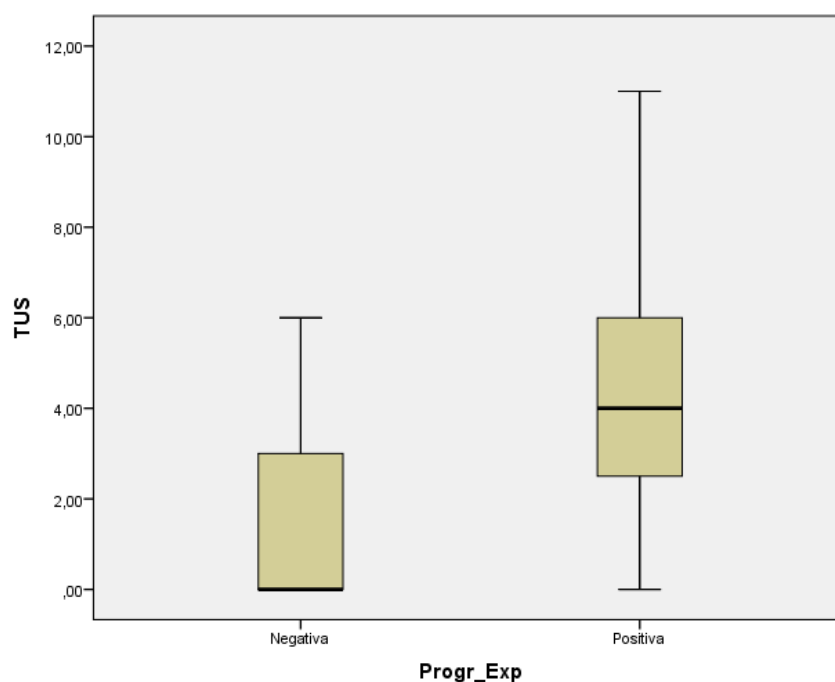
a. Corrección de la significación de Lilliefors

La variable experiencia en programación no supera el test de normalidad.

1.18.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Progr. Exp				Estadístico	Error típ.
TUS	Negativa	Media		1,5000	1,50000
		Intervalo de confianza para la media al 95%	Límite inferior	-3,2737	
			Límite superior	6,2737	
		Media recortada al 5%		1,3333	
		Mediana		,0000	
		Varianza		9,000	
		Desv. típ.		3,00000	
		Mínimo		,00	
		Máximo		6,00	
		Rango		6,00	
		Amplitud intercuartil		4,50	
		Asimetría		2,000	1,014
		Curtosis		4,000	2,619
	Positiva	Media		3,9500	,61334
		Intervalo de confianza para la media al 95%	Límite inferior	2,6663	
			Límite superior	5,2337	
		Media recortada al 5%		3,7778	
		Mediana		4,0000	
		Varianza		7,524	
		Desv. típ.		2,74293	
		Mínimo		,00	
		Máximo		11,00	
		Rango		11,00	
		Amplitud intercuartil		3,75	
		Asimetría		,398	,512
		Curtosis		,995	,992

1.18.3 DIAGRAMA DE CAJAS: TUS



1.18.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Progr_Exp.	Prueba U de Mann-Whitney de muestras independientes	,183 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La experiencia de programación influye en la variable respuesta: TUS.

1.19 TUS – EXPERIENCIA EN TDD

1.19.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	TDD_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,313	12	,002	,736	12	,002
	Positiva	,213	12	,140	,894	12	,132

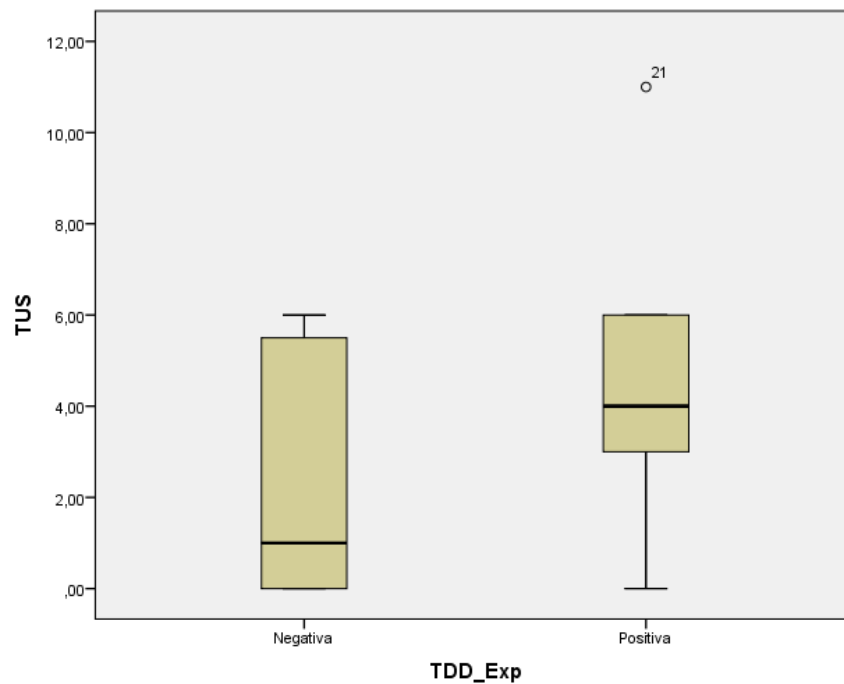
a. Corrección de la significación de Lilliefors

La variable experiencia en TDD no supera el test de normalidad.

1.19.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				Estadístico	Error típ.
TDD_Exp					
TUS	Negativa	Media		2,5000	,81184
		Intervalo de confianza para la media al 95%	Límite inferior	,7131	
			Límite superior	4,2869	
		Media recortada al 5%		2,4444	
		Mediana		1,0000	
		Varianza		7,909	
		Desv. típ.		2,81231	
		Mínimo		,00	
		Máximo		6,00	
		Rango		6,00	
		Amplitud intercuartil		5,75	
		Asimetría		,324	,637
		Curtosis		-2,112	1,232
	Positiva	Media		4,5833	,76335
		Intervalo de confianza para la media al 95%	Límite inferior	2,9032	
			Límite superior	6,2635	
		Media recortada al 5%		4,4815	
		Mediana		4,0000	
		Varianza		6,992	
		Desv. típ.		2,64432	
		Mínimo		,00	
		Máximo		11,00	
		Rango		11,00	
		Amplitud intercuartil		3,00	
		Asimetría		,968	,637
		Curtosis		2,953	1,232

1.19.3 DIAGRAMA DE CAJAS: TUS



1.19.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de TDD_Exp.	Prueba U de Mann-Whitney de muestras independientes	,160 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La experiencia en TDD influye en la variable respuesta: TUS

1.20 TUS – EXPERIENCIA EN JUNIT

1.20.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Junit_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,355	9	,002	,687	9	,001
	Positiva	,178	15	,200 [*]	,917	15	,171

*. Este es un límite inferior de la significación verdadera.

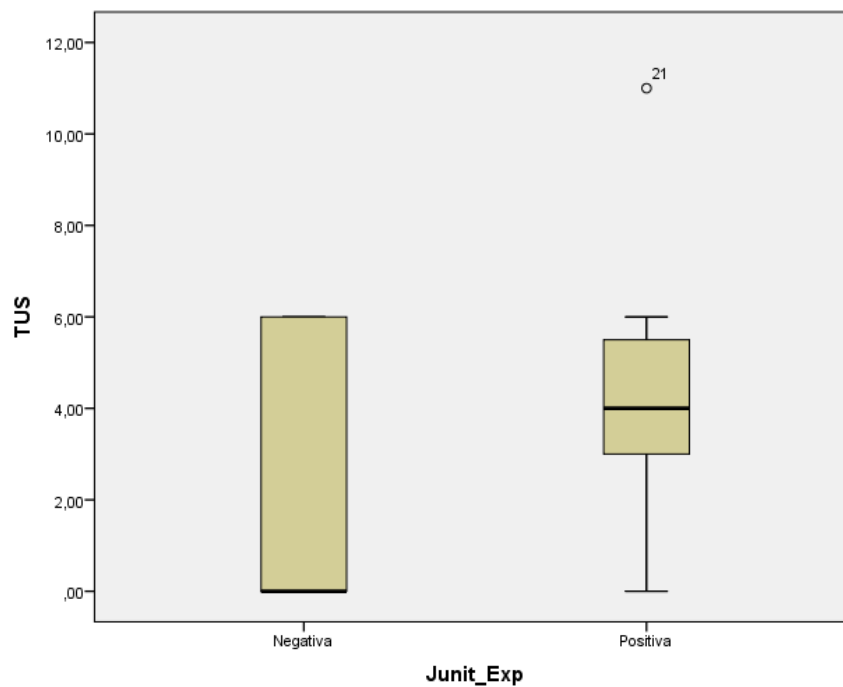
a. Corrección de la significación de Lilliefors

La variable experiencia en JUnit no supera el test de normalidad.

1.20.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Junit_Exp			Estadístico	Error típ.	
TUS	Negativa	Media	2,5556	1,01531	
		Intervalo de confianza para la media al 95%	Límite inferior	,2142	
			Límite superior	4,8969	
		Media recortada al 5%	2,5062		
		Mediana	,0000		
		Varianza	9,278		
		Desv. típ.	3,04594		
		Mínimo	,00		
		Máximo	6,00		
		Rango	6,00		
		Amplitud intercuartil	6,00		
		Asimetría	,306	,717	
		Curtosis	-2,483	1,400	
	Positiva	Media	4,1333	,69602	
		Intervalo de confianza para la media al 95%	Límite inferior	2,6405	
			Límite superior	5,6261	
		Media recortada al 5%	3,9815		
		Mediana	4,0000		
		Varianza	7,267		
		Desv. típ.	2,69568		
		Mínimo	,00		
		Máximo	11,00		
		Rango	11,00		
		Amplitud intercuartil	3,00		
		Asimetría	,796	,580	
		Curtosis	2,204	1,121	

1.20.3 DIAGRAMA DE CAJAS: TUS



1.20.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Junit_Exp.	Prueba U de Mann-Whitney de muestras independientes	,379 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La experiencia en el uso de JUnit es influyente en la variable respuesta: TUS-

1.21 TUS – EXPERIENCIA IDE

1.21.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	IDE_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,354	4	.	,856	4	,247
	Positiva	,211	20	,020	,804	20	,001

a. Corrección de la significación de Lilliefors

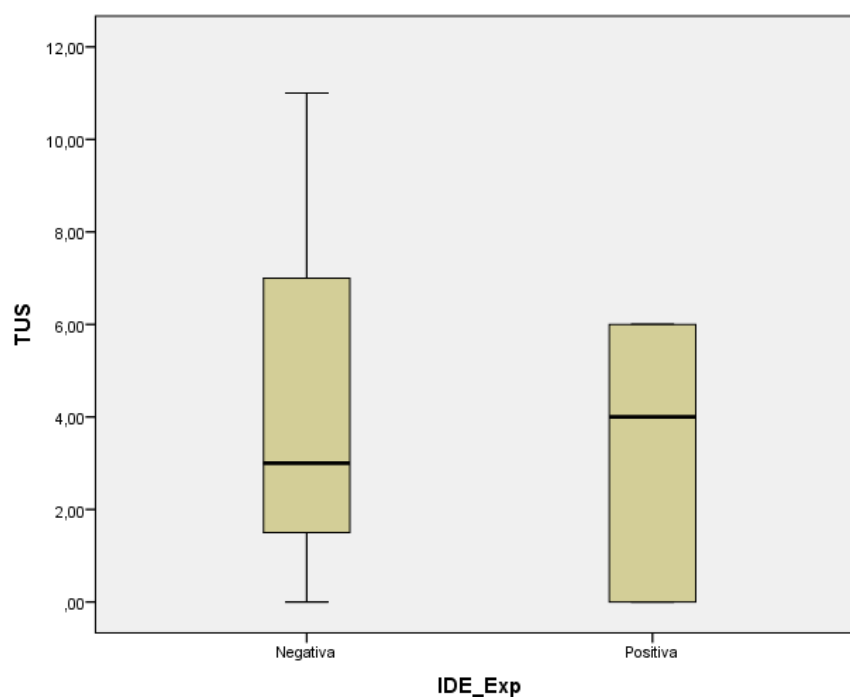
La variable experiencia en el uso de IDEs no supera el test de normalidad.

1.21.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

IDE_Exp			Estadístico	Error tip.
TUS	Negativa	Media	4,2500	2,35850
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	-3,2558 11,7558
		Media recortada al 5%	4,1111	
		Mediana	3,0000	
		Varianza	22,250	
		Desv. típ.	4,71699	
		Mínimo	,00	
		Máximo	11,00	
		Rango	11,00	
		Amplitud intercuartil	8,25	
		Asimetría	1,441	1,014
		Curtosis	2,707	2,619
	Positiva	Media	3,4000	,56382
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	2,2199 4,5801
		Media recortada al 5%	3,4444	
		Mediana	4,0000	
		Varianza	6,358	
		Desv. típ.	2,52149	
		Mínimo	,00	
		Máximo	6,00	
		Rango	6,00	
		Amplitud intercuartil	6,00	
		Asimetría	-,439	,512
		Curtosis	-1,574	,992

1.21.3 IAGRAMA DE CAJAS: TUS



1.21.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de IDE_Exp.	Prueba U de Mann-Whitney de muestras independientes	,911 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La Experiencia en el uso de IDEs no resulta tener una diferencia significativa en la media, con lo cual parece ser que no influye en la variable respuesta: TUS. Esta variable tampoco presenta influencia en las variables respuesta: Productividad y Calidad.

1.22 TUS – EXPERIENCIA PRUEBAS DE UNIDAD

1.22.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	MetDes_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,230	5	,200*	,901	5	,413
	Positiva	,183	19	,093	,853	19	,007

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

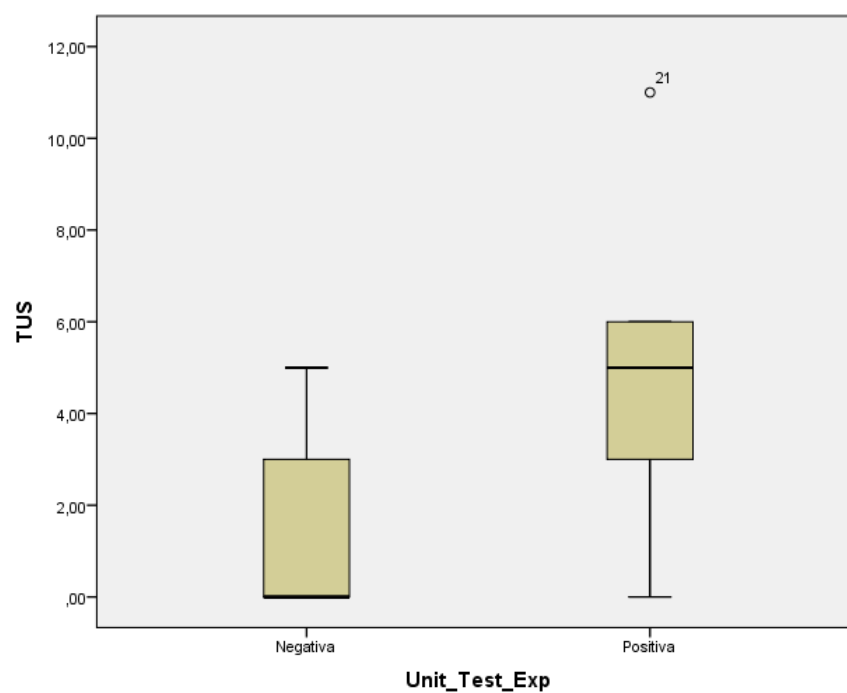
La variable experiencia en pruebas de unidad no supera el test de normalidad.

1.22.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

			Estadístico	Error típ.
TUS	Media		3,5417	,58662
	Intervalo de confianza para la media al 95%	Límite inferior	2,3281	
		Límite superior	4,7552	
	Media recortada al 5%		3,3704	
	Mediana		4,0000	
	Varianza		8,259	
	Desv. típ.		2,87386	
	Mínimo		,00	
	Máximo		11,00	
	Rango		11,00	
	Amplitud intercuartil		6,00	
	Asimetría		,397	,472
	Curtosis		,236	,918

1.22.3 DIAGRAMA DE CAJAS: TUS



1.22.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Unit_Test_Exp.	Prueba U de Mann-Whitney de muestras independientes	,007 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La experiencia en pruebas de unidad resulta influyente en la variable respuesta: TUS.

1.23 TUS – EXPERIENCIA DE PROGRAMACIÓN

1.23.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Programacion_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,293	4	.	,860	4	,262
	Positiva	,180	20	,090	,890	20	,027

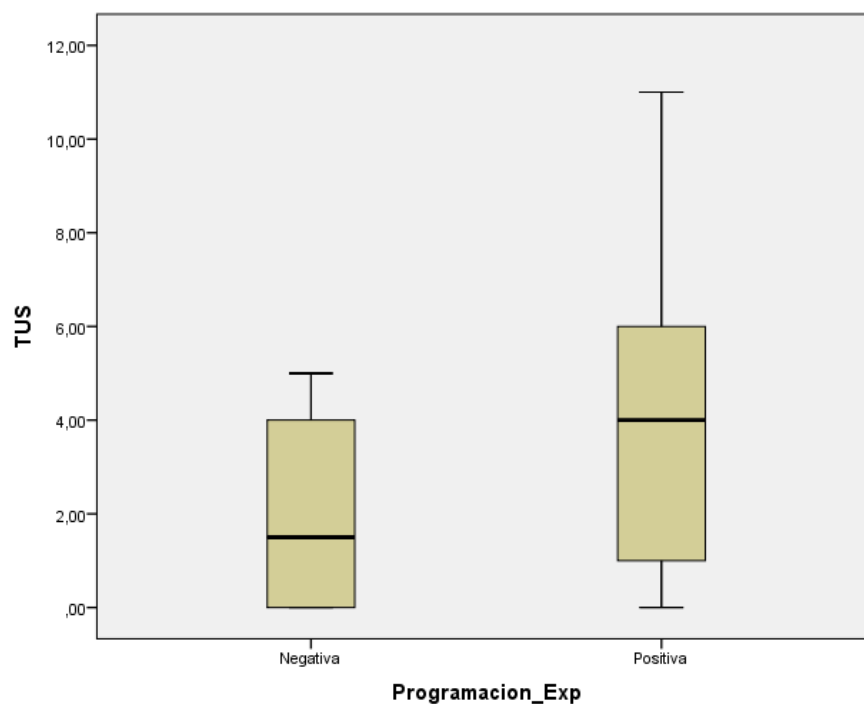
a. Corrección de la significación de Lilliefors

La variable Experiencia de programación no supera el test de normalidad.

1.23.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Programacion Exp			Estadístico	Error típ.	
TUS	Negativa	Media	2,0000	1,22474	
		Intervalo de confianza para la media al 95%	Límite inferior	-1,8977	
			Límite superior	5,8977	
		Media recortada al 5%	1,9444		
		Mediana	1,5000		
		Varianza	6,000		
		Desv. típ.	2,44949		
		Mínimo	,00		
		Máximo	5,00		
		Rango	5,00		
		Amplitud intercuartil	4,50		
		Asimetría	,544	1,014	
		Curtosis	-2,944	2,619	
	Positiva	Media	3,8500	,65000	
		Intervalo de confianza para la media al 95%	Límite inferior	2,4895	
			Límite superior	5,2105	
		Media recortada al 5%	3,6667		
		Mediana	4,0000		
		Varianza	8,450		
		Desv. típ.	2,90689		
		Mínimo	,00		
		Máximo	11,00		
		Rango	11,00		
		Amplitud intercuartil	5,50		
		Asimetría	,322	,512	
		Curtosis	,362	,992	

1.23.3 DIAGRAMA DE CAJAS



1.23.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Programacion_Exp.	Prueba U de Mann-Whitney de muestras independientes	,241 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

1.24 TUS – METODOLOGÍAS DE DESARROLLO

Prueba de normalidad

Pruebas de normalidad

	MetDes_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,230	5	,200 [*]	,901	5	,413
	Positiva	,183	19	,093	,853	19	,007

*. Este es un límite inferior de la significación verdadera.

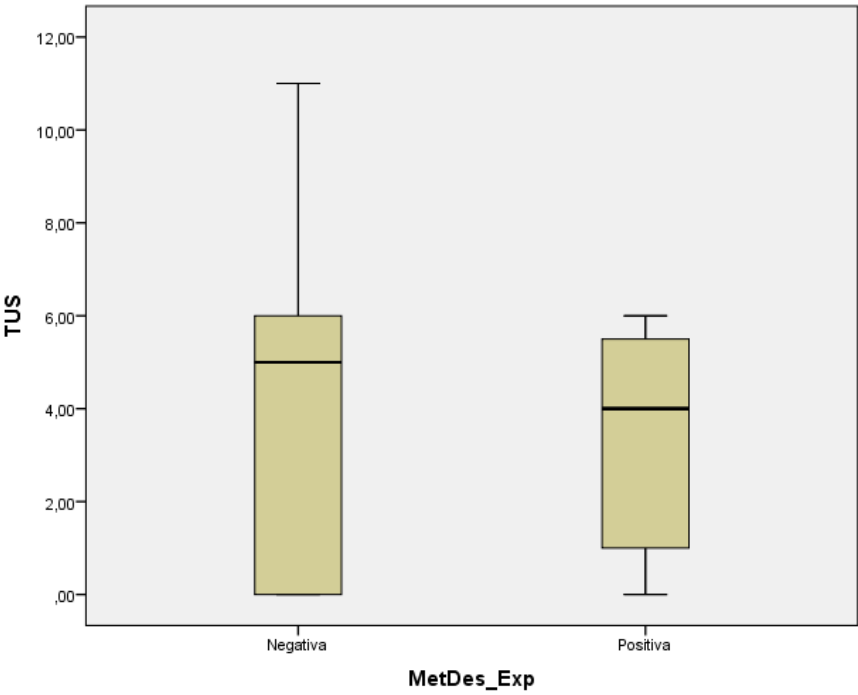
a. Corrección de la significación de Lilliefors

La variable Metodologías de desarrollo no supera el test de normalidad.

1.24.1 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
MetDes_Exp				Estadístico	Error típ.
TUS	Negativa	Media		4,4000	2,06398
		Intervalo de confianza para la media al 95%	Límite inferior	-1,3305	
			Límite superior	10,1305	
		Media recortada al 5%		4,2778	
		Mediana		5,0000	
		Varianza		21,300	
		Desv. típ.		4,61519	
		Mínimo		,00	
		Máximo		11,00	
		Rango		11,00	
		Amplitud intercuartil		8,50	
		Asimetría		,515	,913
		Curtosis		-,688	2,000
	Positiva	Media		3,3158	,54102
		Intervalo de confianza para la media al 95%	Límite inferior	2,1791	
			Límite superior	4,4524	
		Media recortada al 5%		3,3509	
		Mediana		4,0000	
		Varianza		5,561	
		Desv. típ.		2,35826	
		Mínimo		,00	
		Máximo		6,00	
		Rango		6,00	
		Amplitud intercuartil		6,00	
		Asimetría		-,367	,524
		Curtosis		-1,356	1,014

1.24.2 DIAGRAMA DE CAJAS



1.24.3 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de MetDes_Exp.	Prueba U de Mann-Whitney de muestras independientes	,679 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

2 ANÁLISIS DICOTÓMICO BOWLING SCOREKEEPER

2.1 PRODUCTIVIDAD – EXPERIENCIA ACADÉMICA

2.1.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Académic_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,298	4	.	,849	4	,224
	Positiva	,240	20	,004	,800	20	,001

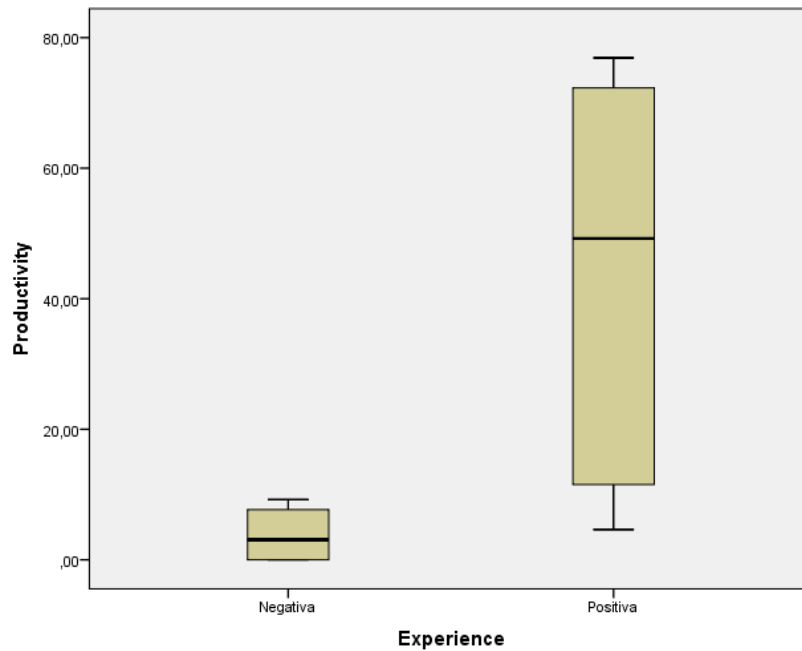
a. Corrección de la significación de Lilliefors

2.1.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Académic Exp				Estadístico	Error típ.
Productivity	Negativa	Media		3,8450	2,30722
		Intervalo de confianza para la media al 95%	Límite inferior	-3,4976	
			Límite superior	11,1876	
		Media recortada al 5%		3,7594	
		Mediana		3,0750	
		Varianza		21,293	
		Desv. típ.		4,61444	
		Mínimo		,00	
		Máximo		9,23	
		Rango		9,23	
		Amplitud intercuartil		8,46	
		Asimetría		,371	1,014
		Curtosis		-3,896	2,619
	Positiva	Media		43,4615	6,72526
		Intervalo de confianza para la media al 95%	Límite inferior	29,3854	
			Límite superior	57,5376	
		Media recortada al 5%		43,7606	
		Mediana		49,2300	
		Varianza		904,582	
		Desv. típ.		30,07626	
		Mínimo		4,62	
		Máximo		76,92	
		Rango		72,30	
		Amplitud intercuartil		61,16	
		Asimetría		-,121	,512
		Curtosis		-1,992	,992

2.1.3 DIAGRAMA DE CAJAS: TUS



2.1.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Académic_Exp.	Prueba U de Mann-Whitney de muestras independientes	,002 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Según la gráfica no cabe duda de que la variable experiencia académica influye en la variable respuesta: Productividad.

2.2 PRODUCTIVIDAD – EXPERIENCIA JAVA

2.2.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	,227	24	,002	,801	24	,000

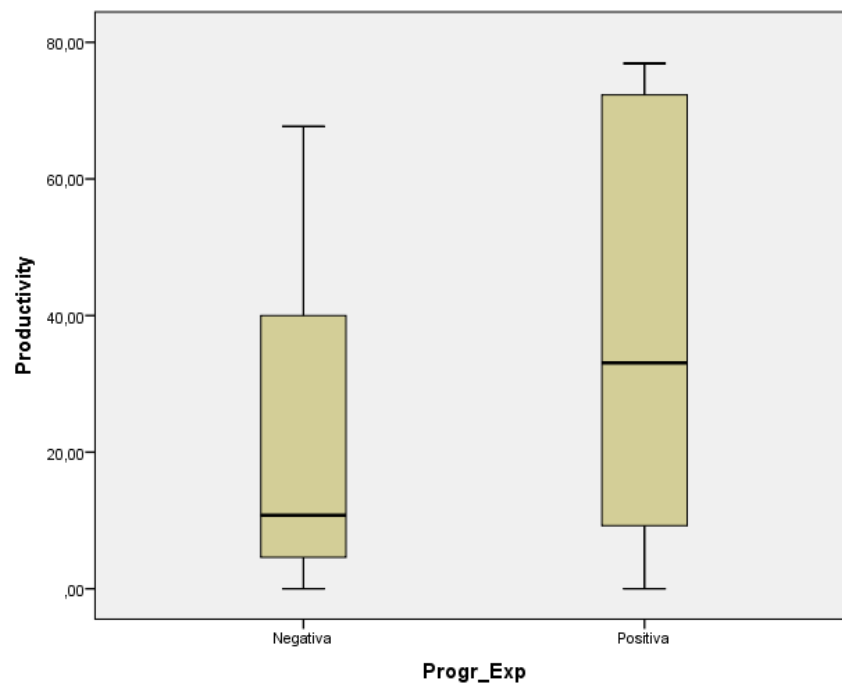
a. Corrección de la significación de Lilliefors

La variable experiencia de programación no supera el test de normalidad.

2.2.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos			Estadístico	Error típ.
Productivity	Media		36,8588	6,38194
	Intervalo de confianza para la media al 95%	Límite inferior	23,6567	
		Límite superior	50,0608	
	Media recortada al 5%		36,6808	
	Mediana		23,8450	
	Varianza		977,499	
	Desv. típ.		31,26499	
	Mínimo		,00	
	Máximo		76,92	
	Rango		76,92	
	Amplitud intercuartil		63,85	
	Asimetría		,217	,472
	Curtosis		-1,914	,918

2.2.3 DIAGRAMA DE CAJAS: TUS



2.2.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Progr_Exp.	Prueba U de Mann-Whitney de muestras independientes	,241 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La variable experiencia de programación influye sobre la variable respuesta: Productividad.

2.3 PRODUCTIVIDAD – EXPERIENCIA EN TDD

2.3.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	TDD_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,230	12	,078	,826	12	,019
	Positiva	,241	12	,053	,770	12	,004

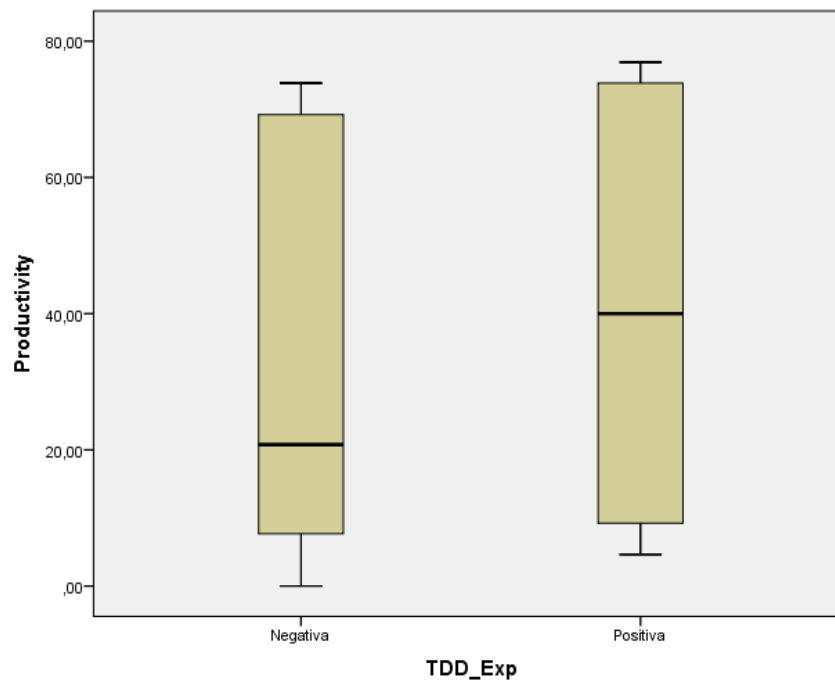
a. Corrección de la significación de Lilliefors

La variable Experiencia en TDD no supera el test de normalidad.

2.3.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
TDD_Exp				Estadístico	Error típ.
Productivity	Negativa	Media		32,5642	8,80105
		Intervalo de confianza para la media al 95%	Límite inferior	13,1932	
			Límite superior	51,9352	
		Media recortada al 5%		32,0796	
		Mediana		20,7700	
		Varianza		929,502	
		Desv. típ.		30,48774	
		Mínimo		,00	
		Máximo		73,85	
		Rango		73,85	
		Amplitud intercuartil		63,08	
		Asimetría		,444	,637
		Curtosis		-1,765	1,232
	Positiva	Media		41,1533	9,46097
		Intervalo de confianza para la media al 95%	Límite inferior	20,3299	
			Límite superior	61,9768	
		Media recortada al 5%		41,1959	
		Mediana		40,0000	
		Varianza		1074,120	
		Desv. típ.		32,77376	
		Mínimo		4,62	
		Máximo		76,92	
		Rango		72,30	
		Amplitud intercuartil		66,15	
		Asimetría		,012	,637
		Curtosis		-2,290	1,232

2.3.3 DIAGRAMA DE CAJAS: TUS



2.3.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de TDD_Exp.	Prueba U de Mann-Whitney de muestras independientes	,410 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La variable experiencia en TDD influye sobre la variable respuesta: Productividad.

2.4 PRODUCTIVIDAD – EXPERIENCIA EN JUNIT

2.4.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Junit_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,360	9	,001	,734	9	,004
	Positiva	,233	15	,028	,838	15	,012

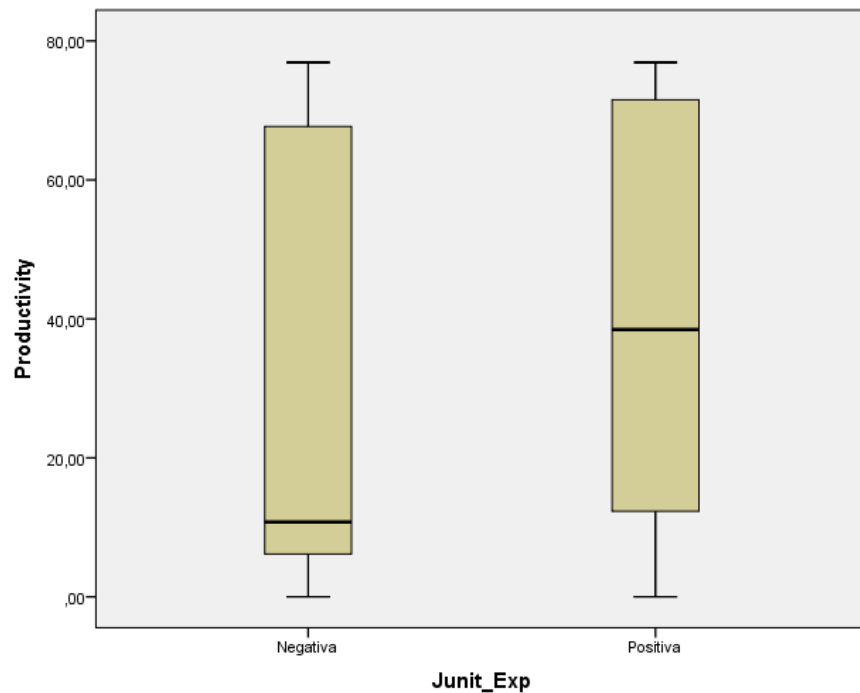
a. Corrección de la significación de Lilliefors

La variable Experiencia en IUnit no supera el test de normalidad.

2.4.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				Estadístico	Error típ.
Junit_Exp					
Productivity	Negativa	Media		29,0600	11,03277
		Intervalo de confianza para la media al 95%	Límite inferior	3,6184	
			Límite superior	54,5016	
		Media recortada al 5%		28,0156	
		Mediana		10,7700	
		Varianza		1095,498	
		Desv. típ.		33,09830	
		Mínimo		,00	
		Máximo		76,92	
		Rango		76,92	
		Amplitud intercuartil		65,38	
		Asimetría		,830	,717
		Curtosis		-1,628	1,400
	Positiva	Media		41,5380	7,82021
		Intervalo de confianza para la media al 95%	Límite inferior	24,7653	
			Límite superior	58,3107	
		Media recortada al 5%		41,8800	
		Mediana		38,4600	
		Varianza		917,335	
		Desv. típ.		30,28754	
		Mínimo		,00	
		Máximo		76,92	
		Rango		76,92	
		Amplitud intercuartil		61,54	
		Asimetría		-,060	,580
		Curtosis		-1,962	1,121

2.4.3 DIAGRAMA DE CAJAS: TUS



2.4.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Junit_Exp.	Prueba U de Mann-Whitney de muestras independientes	,318 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Dada la diferencia de medias se considera que la variable experiencia en JUnit influye en la variable respuesta: Productividad

2.5 PRODUCTIVIDAD – EXPERIENCIA IDE

2.5.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	IDE_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,400	4	.	,716	4	,018
	Positiva	,276	20	,000	,797	20	,001

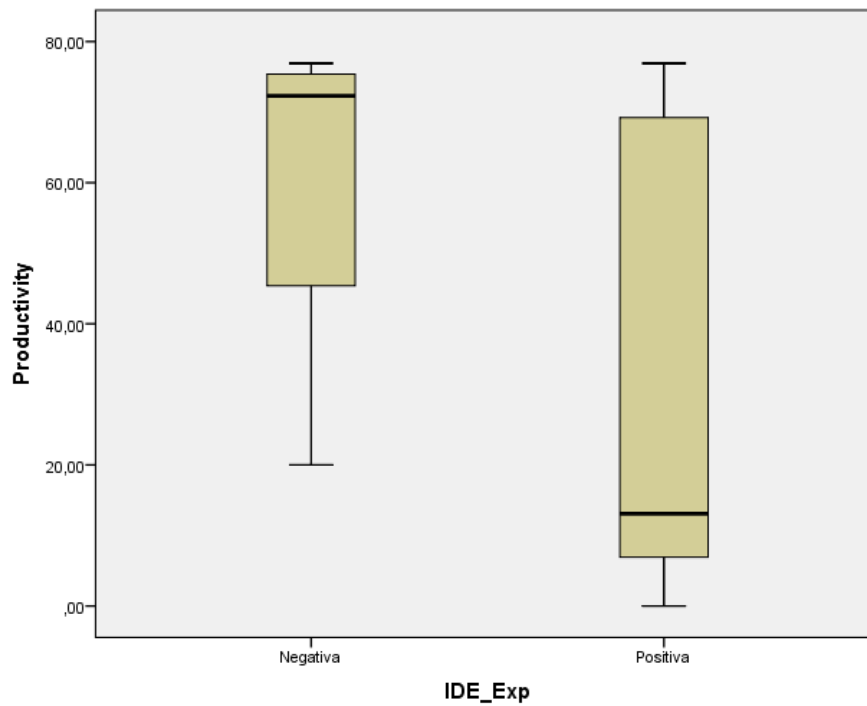
a. Corrección de la significación de Lilliefors

La variable Experiencia en el uso de IDEs no supera el test de normalidad.

2.5.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
IDE_Exp				Estadístico	Error típ.
Productivity	Negativa	Media		60,3850	13,52007
		Intervalo de confianza para la media al 95%	Límite inferior	17,3581	
			Límite superior	103,4119	
		Media recortada al 5%		61,7100	
		Mediana		72,3100	
		Varianza		731,170	
		Desv. típ.		27,04015	
		Mínimo		20,00	
		Máximo		76,92	
		Rango		56,92	
		Amplitud intercuartil		43,46	
		Asimetría		-1,948	1,014
		Curtosis		3,829	2,619
	Positiva	Media		32,1535	6,81180
		Intervalo de confianza para la media al 95%	Límite inferior	17,8962	
			Límite superior	46,4108	
		Media recortada al 5%		31,4528	
		Mediana		13,0800	
		Varianza		928,013	
		Desv. típ.		30,46331	
		Mínimo		,00	
		Máximo		76,92	
		Rango		76,92	
		Amplitud intercuartil		63,47	
		Asimetría		,500	,512
		Curtosis		-1,687	,992

2.5.3 DIAGRAMA DE CAJAS: TUS



2.5.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de IDE_Exp.	Prueba U de Mann-Whitney de muestras independientes	,068 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

El resultado obtenido resulta curioso, los sujetos con menos experiencia en el uso de IDEs presentan más productividad que lo que no la tiene. Este resultado hay que analizarlo más detenidamente.

2.6 PRODUCTIVIDAD – EXPERIENCIA PRUEBAS DE UNIDAD

2.6.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Unit_Test_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,392	8	,001	,678	8	,001
	Positiva	,224	16	,031	,823	16	,006

a. Corrección de la significación de Lilliefors

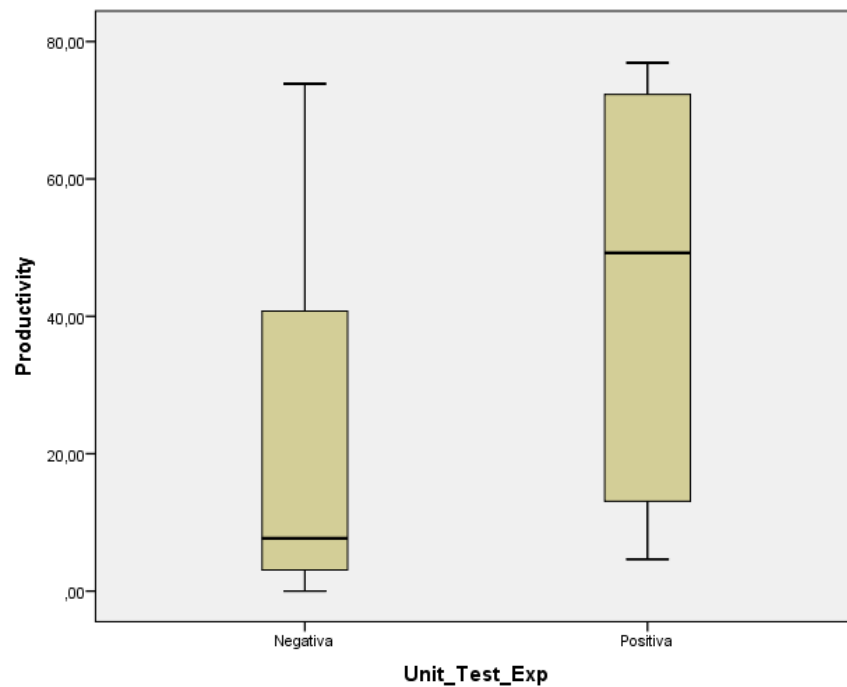
La variable experiencia en pruebas de unidad no supera el test de normalidad.

2.6.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Unit Test Exp				Estadístico	Error típ.
Productivity	Negativa	Media		22,1150	11,04088
		Intervalo de confianza para la media al 95%	Límite inferior	-3,9925	
			Límite superior	48,2225	
		Media recortada al 5%		20,4694	
		Mediana		7,6900	
		Varianza		975,209	
		Desv. típ.		31,22833	
		Mínimo		,00	
		Máximo		73,85	
		Rango		73,85	
		Amplitud intercuartil		54,23	
		Asimetría		1,378	,752
		Curtosis		-,062	1,481
	Positiva	Media		44,2306	7,37323
		Intervalo de confianza para la media al 95%	Límite inferior	28,5150	
			Límite superior	59,9463	
		Media recortada al 5%		44,6151	
		Mediana		49,2300	
		Varianza		869,832	
		Desv. típ.		29,49292	
		Mínimo		4,62	
		Máximo		76,92	
		Rango		72,30	
		Amplitud intercuartil		59,62	
		Asimetría		-,143	,564
		Curtosis		-1,979	1,091

2.6.3 DIAGRAMA DE CAJAS: TUS



2.6.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Unit_Test_Exp.	Prueba U de Mann-Whitney de muestras independientes	,038 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La variable experiencia en pruebas de unidad influye sobre la variable respuesta: Productividad.

2.7 PRODUCTIVIDAD – EXPERIENCIA DE PROGRAMACIÓN

2.7.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Programacion_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,290	4	.	,863	4	,271
	Positiva	,240	20	,004	,799	20	,001

a. Corrección de la significación de Lilliefors

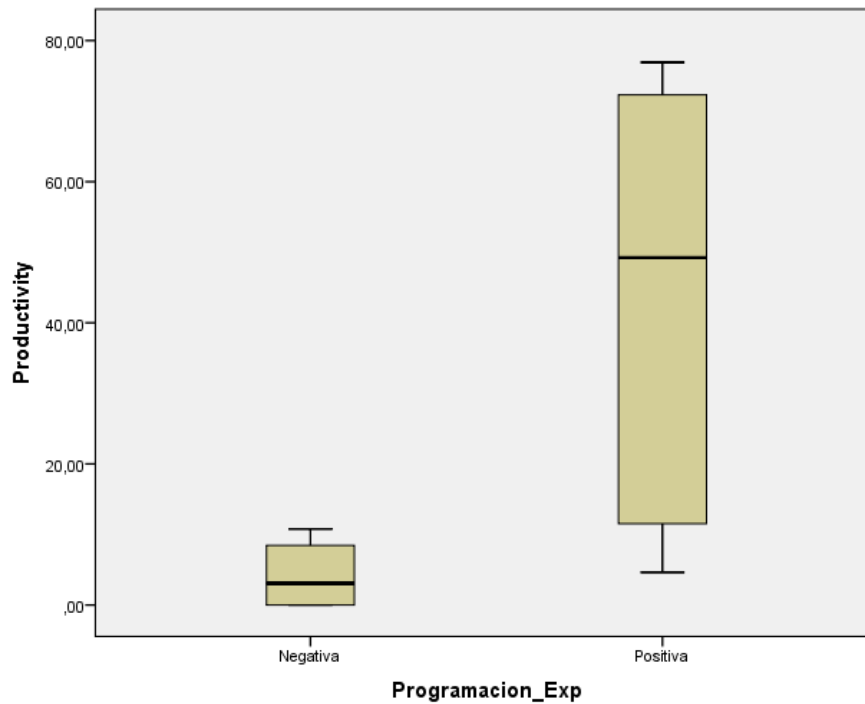
La variable Experiencia de programación supera el test de normalidad.

2.7.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Programacion_Exp				Estadístico	Error típ.
Productivity	Negativa	Media		4,2300	2,61795
		Intervalo de confianza para la media al 95%	Límite inferior	-4,1015	
			Límite superior	12,5615	
		Media recortada al 5%		4,1017	
		Mediana		3,0750	
		Varianza		27,415	
		Desv. típ.		5,23590	
		Mínimo		,00	
		Máximo		10,77	
		Rango		10,77	
		Amplitud intercuartil		9,62	
		Asimetría		,629	1,014
		Curtosis		-2,486	2,619
	Positiva	Media		43,3845	6,74537
		Intervalo de confianza para la media al 95%	Límite inferior	29,2663	
			Límite superior	57,5027	
		Media recortada al 5%		43,6750	
		Mediana		49,2300	
		Varianza		910,000	
		Desv. típ.		30,16620	
		Mínimo		4,62	
		Máximo		76,92	
		Rango		72,30	
		Amplitud intercuartil		61,16	
		Asimetría		-,123	,512
		Curtosis		-1,992	,992

2.7.3 DIAGRAMA DE CAJAS



2.7.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de Programacion_Exp.	Prueba U de Mann-Whitney de muestras independientes	,005 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

2.8 PRODUCTIVIDAD – METODOLOGÍAS DE DESARROLLO

2.8.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	MetDes_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,423	5	,004	,652	5	,003
	Positiva	,229	19	,010	,818	19	,002

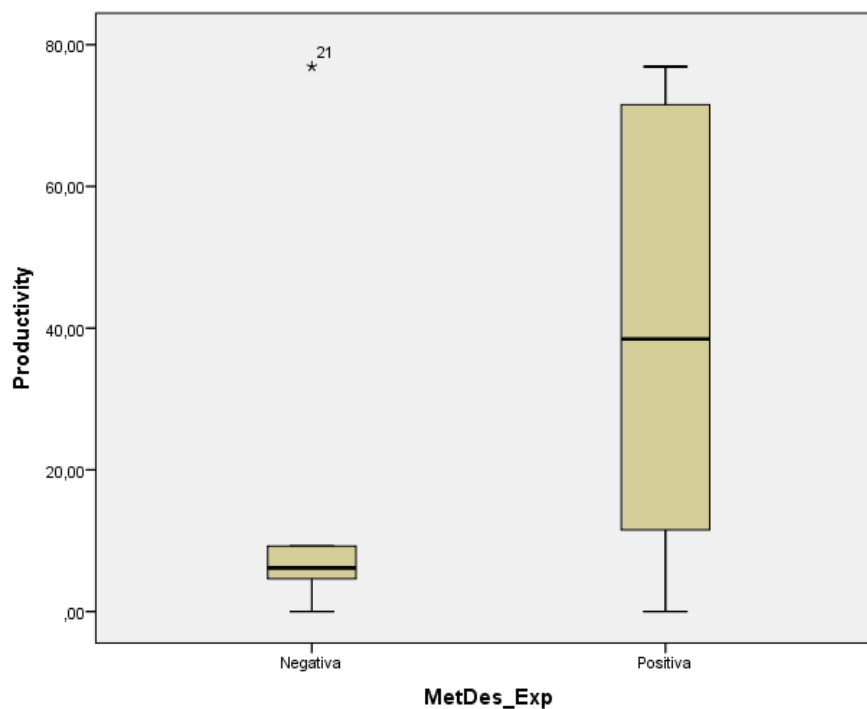
a. Corrección de la significación de Lilliefors

La variable Experiencia en metodologías de desarrollo no supera el test de normalidad.

2.8.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				Estadístico	Error típ.
Productivity	MetDes	Exp			
Productivity	Negativa	Media		19,3840	14,46089
		Intervalo de confianza para la media al 95%	Límite inferior	-20,7659	
			Límite superior	59,5339	
		Media recortada al 5%		17,2644	
		Mediana		6,1500	
		Varianza		1045,587	
		Desv. típ.		32,33554	
		Mínimo		,00	
		Máximo		76,92	
		Rango		76,92	
		Amplitud intercuartil		40,77	
		Asimetría		2,176	,913
		Curtosis		4,793	2,000
	Positiva	Media		41,4574	6,91881
		Intervalo de confianza para la media al 95%	Límite inferior	26,9215	
			Límite superior	55,9932	
		Media recortada al 5%		41,7904	
		Mediana		38,4600	
		Varianza		909,528	
		Desv. típ.		30,15838	
		Mínimo		,00	
		Máximo		76,92	
		Rango		76,92	
		Amplitud intercuartil		61,54	
		Asimetría		-,042	,524
		Curtosis		-1,977	1,014

2.8.3 DIAGRAMA DE CAJAS



2.8.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Productivity es la misma entre las categorías de MetDes_Exp.	Prueba U de Mann-Whitney de muestras independientes	,103 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

2.9 CALIDAD – EXPERIENCIA ACADÉMICA

2.9.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Académic_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,307	4	.	,729	4	,024
	Positiva	,153	20	,200 [*]	,964	20	,618

*. Este es un límite inferior de la significación verdadera.

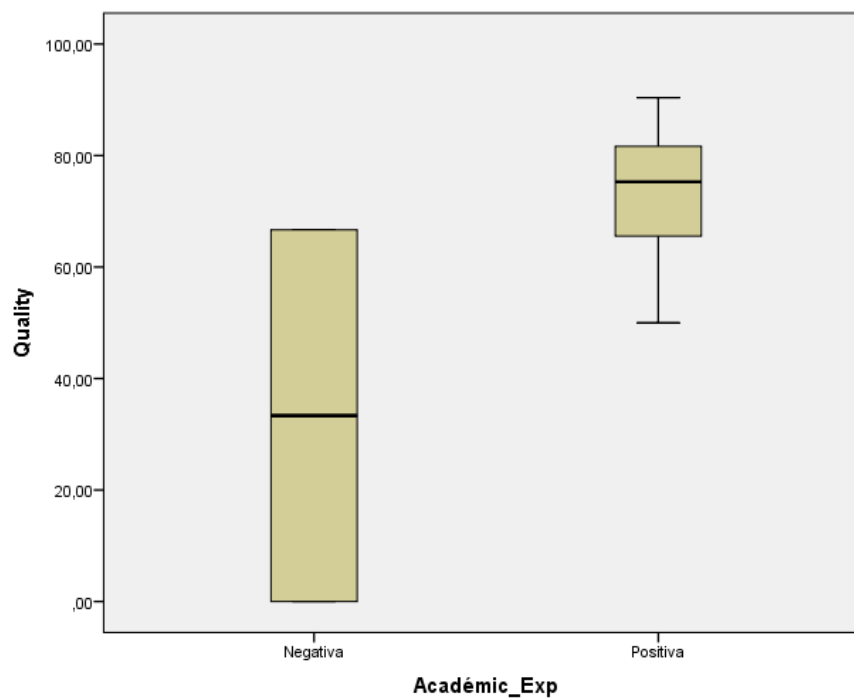
a. Corrección de la significación de Lilliefors

La variable Experiencia académica no supera el test de normalidad.

2.9.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Académic Exp			Estadístico	Error típ.
Quality	Negativa	Media	33,3350	19,24597
		Intervalo de confianza para la media al 95%	Límite inferior	-27,9143
			Límite superior	94,5843
		Media recortada al 5%	33,3350	
		Mediana	33,3350	
		Varianza	1481,630	
		Desv. típ.	38,49194	
		Mínimo	,00	
		Máximo	66,67	
		Rango	66,67	
		Amplitud intercuartil	66,67	
		Asimetría	,000	1,014
		Curtosis	-6,000	2,619
	Positiva	Media	73,2940	2,39972
		Intervalo de confianza para la media al 95%	Límite inferior	68,2713
			Límite superior	78,3167
		Media recortada al 5%	73,6394	
		Mediana	75,3050	
		Varianza	115,173	
		Desv. típ.	10,73186	
		Mínimo	50,00	
		Máximo	90,37	
		Rango	40,37	
		Amplitud intercuartil	16,67	
		Asimetría	-,489	,512
		Curtosis	-,252	,992

2.9.3 DIAGRAMA DE CAJAS: TUS



2.9.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de Académic_Exp.	Prueba U de Mann-Whitney de muestras independientes	,023 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Según la grafica la experiencia académica influye sobre la variable respuesta: Calidad.

2.10 CALIDAD – EXPERIENCIA JAVA

2.10.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Progr_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,318	4	.	,813	4	,128
	Positiva	,193	20	,050	,742	20	,000

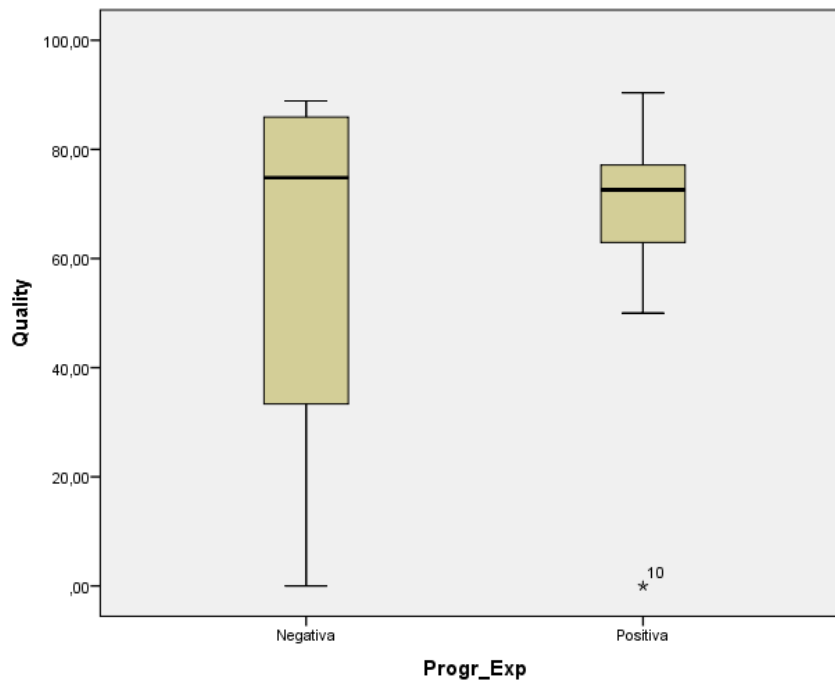
a. Corrección de la significación de Lilliefors

La variable Experiencia de programación no supera el test de normalidad.

2.10.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Progr. Exp			Estadístico	Error típ.
Quality	Negativa	Media	59,6300	20,42412
		Intervalo de confianza para la media al 95%	Límite inferior	-5,3687
			Límite superior	124,6287
		Media recortada al 5%	61,3172	
		Mediana	74,8150	
		Varianza	1668,578	
		Desv. típ.	40,84824	
		Mínimo	,00	
		Máximo	88,89	
		Rango	88,89	
		Amplitud intercuartil	70,74	
		Asimetría	-1,701	1,014
		Curtosis	2,872	2,619
	Positiva	Media	68,0350	4,20350
		Intervalo de confianza para la media al 95%	Límite inferior	59,2370
			Límite superior	76,8330
		Media recortada al 5%	70,5739	
		Mediana	72,5800	
		Varianza	353,388	
		Desv. típ.	18,79861	
		Mínimo	,00	
		Máximo	90,37	
		Rango	90,37	
		Amplitud intercuartil	15,26	
		Asimetría	-2,643	,512
		Curtosis	9,159	,992

2.10.3 DIAGRAMA DE CAJAS: TUS



2.10.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de Progr_Exp.	Prueba U de Mann-Whitney de muestras independientes	,737 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Las medias no tienen una diferencia muy significativa, con lo cual no se puede afirmar a priori que la variable experiencia de programación influye sobre la variable respuesta: Calidad.

2.11 CALIDAD – EXPERIENCIA EN TDD

2.11.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	TDD_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,333	12	,001	,715	12	,001
	Positiva	,131	12	,200 [*]	,945	12	,561

*. Este es un límite inferior de la significación verdadera.

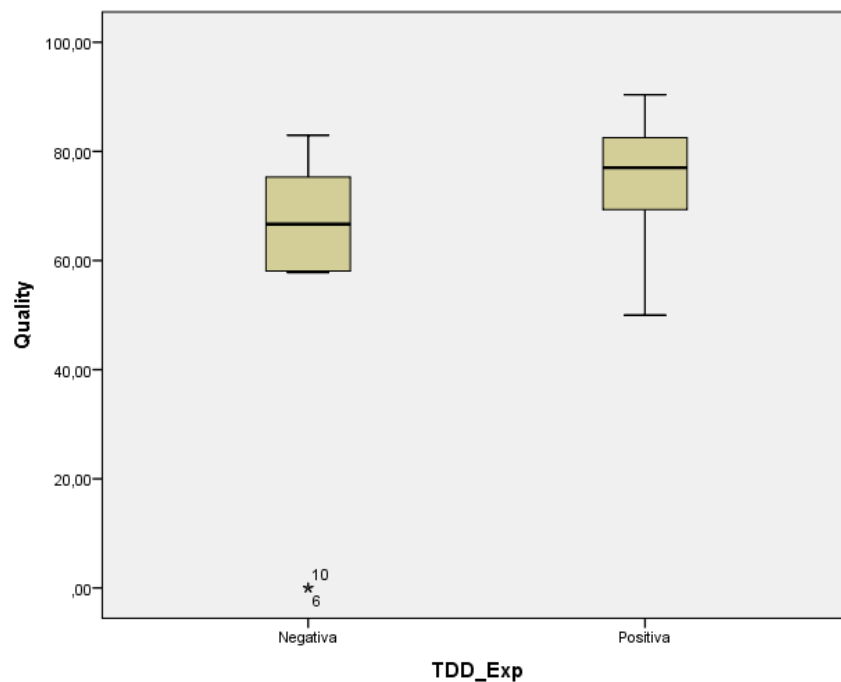
a. Corrección de la significación de Lilliefors

La variable experiencia en TDD no supera el test de normalidad.

2.11.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
TDD_Exp			Estadístico	Error típ.	
Quality	Negativa	Media	57,7975	8,10299	
		Intervalo de confianza para la media al 95%	Límite inferior	39,9629	
			Límite superior	75,6321	
		Media recortada al 5%	59,6106		
		Mediana	66,6700		
		Varianza	787,901		
		Desv. típ.	28,06957		
		Mínimo	,00		
		Máximo	82,96		
		Rango	82,96		
		Amplitud intercuartil	17,53		
		Asimetría	-1,718	,637	
		Curtosis	1,828	1,232	
	Positiva	Media	75,4708	3,27362	
		Intervalo de confianza para la media al 95%	Límite inferior	68,2656	
			Límite superior	82,6760	
		Media recortada al 5%	76,0581		
		Mediana	76,9950		
		Varianza	128,599		
		Desv. típ.	11,34015		
		Mínimo	50,00		
		Máximo	90,37		
		Rango	40,37		
		Amplitud intercuartil	14,90		
		Asimetría	-,891	,637	
		Curtosis	1,021	1,232	

2.11.3 DIAGRAMA DE CAJAS: TUS



2.11.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de TDD_Exp.	Prueba U de Mann-Whitney de muestras independientes	,045 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La variable experiencia de programación no presenta diferencia en medias significativa como para afirmar que resulta influyente.

2.12 CALIDAD – EXPERIENCIA EN JUNIT

2.12.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Junit_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,215	9	,200 [*]	,814	9	,030
	Positiva	,235	15	,025	,685	15	,000

*. Este es un límite inferior de la significación verdadera.

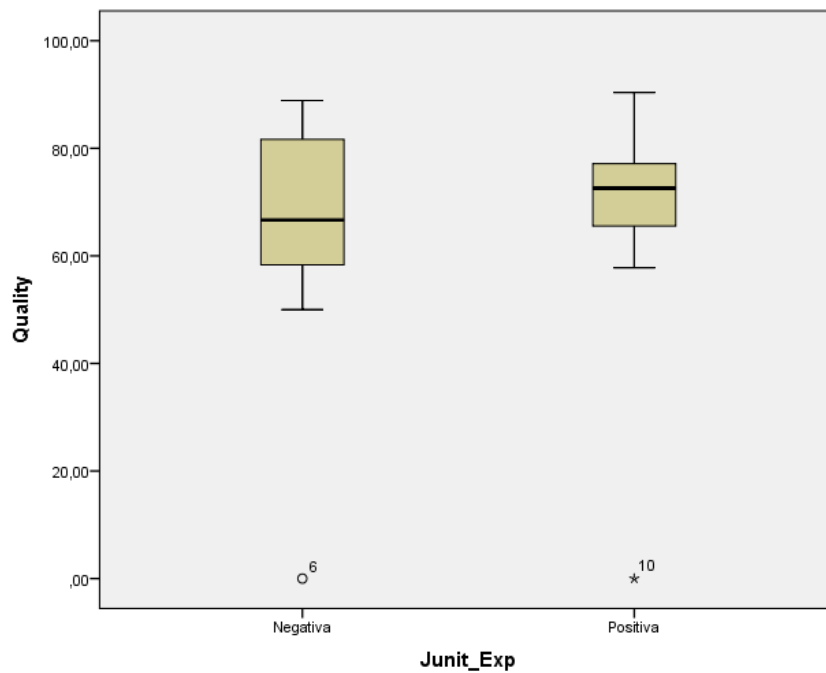
a. Corrección de la significación de Lilliefors

La variable experiencia en JUnit no supera en test de normalidad.

2.12.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Junit_Exp			Estadístico	Error tip.
Quality	Negativa	Media	63,4222	8,94537
		Intervalo de confianza para la media al 95%	Límite inferior	42,7942
			Límite superior	84,0503
		Media recortada al 5%	65,5308	
		Mediana	66,6700	
		Varianza	720,177	
		Desv. típ.	26,83612	
		Mínimo	,00	
		Máximo	88,89	
		Rango	88,89	
		Amplitud intercuartil	28,15	
		Asimetría	-1,877	,717
		Curtosis	4,163	1,400
	Positiva	Media	68,5613	5,36550
		Intervalo de confianza para la media al 95%	Límite inferior	57,0535
			Límite superior	80,0692
		Media recortada al 5%	71,1587	
		Mediana	72,5800	
		Varianza	431,829	
		Desv. típ.	20,78049	
		Mínimo	,00	
		Máximo	90,37	
		Rango	90,37	
		Amplitud intercuartil	13,34	
		Asimetría	-2,810	,580
		Curtosis	9,450	1,121

2.12.3 DIAGRAMA DE CAJAS: TUS



2.12.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de Junit_Exp.	Prueba U de Mann-Whitney de muestras independientes	,726 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La diferencia de medias no resulta significativa como para afirmar que la variable experiencia en Junit influye en la variable respuesta: Calidad.

2.13 CALIDAD – EXPERIENCIA IDE

2.13.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	IDE_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,162	4	.	,996	4	,984
	Positiva	,242	20	,003	,728	20	,000

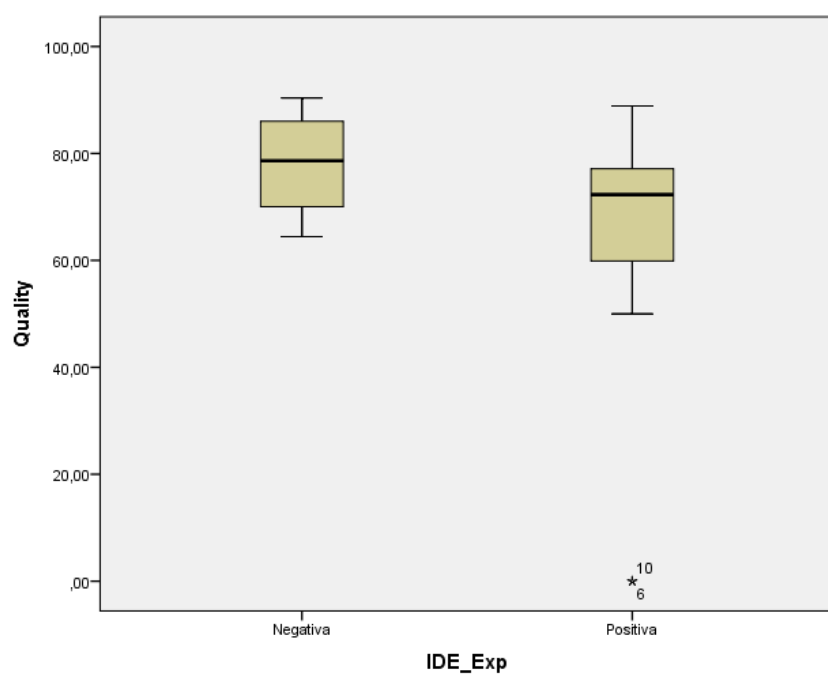
a. Corrección de la significación de Lilliefors

La variable Experiencia en el uso de IDEs no supera en test de normalidad.

2.13.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
IDE_Exp				Estadístico	Error tip.
Quality	Negativa	Media		78,0225	5,44724
		Intervalo de confianza para la media al 95%	Límite inferior	60,6869	
			Límite superior	95,3581	
		Media recortada al 5%		78,0911	
		Mediana		78,6400	
		Varianza		118,690	
		Desv. tip.		10,89449	
		Mínimo		64,44	
		Máximo		90,37	
		Rango		25,93	
		Amplitud intercuartil		20,96	
		Asimetría		-,304	1,014
		Curtosis		,103	2,619
	Positiva	Media		64,3565	5,37464
		Intervalo de confianza para la media al 95%	Límite inferior	53,1072	
			Límite superior	75,6058	
		Media recortada al 5%		66,5689	
		Mediana		72,3100	
		Varianza		577,736	
		Desv. tip.		24,03613	
		Mínimo		,00	
		Máximo		88,89	
		Rango		88,89	
		Amplitud intercuartil		18,35	
		Asimetría		-2,083	,512
		Curtosis		4,069	,992

2.13.3 DIAGRAMA DE CAJAS: TUS



2.13.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de IDE_Exp.	Prueba U de Mann-Whitney de muestras independientes	,241 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Igual que en el apartado anterior la diferencia de las medias no resulta lo suficientemente significativa como para afirmar que existe influencia de la variable Experiencia en el suso de IDEs sobre la variable respuesta: Callidad.

2.14 CALIDAD – EXPERIENCIA PRUEBAS DE UNIDAD

2.14.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Unit_Test_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,336	8	,008	,709	8	,003
	Positiva	,170	16	,200*	,951	16	,511

*. Este es un límite inferior de la significación verdadera.

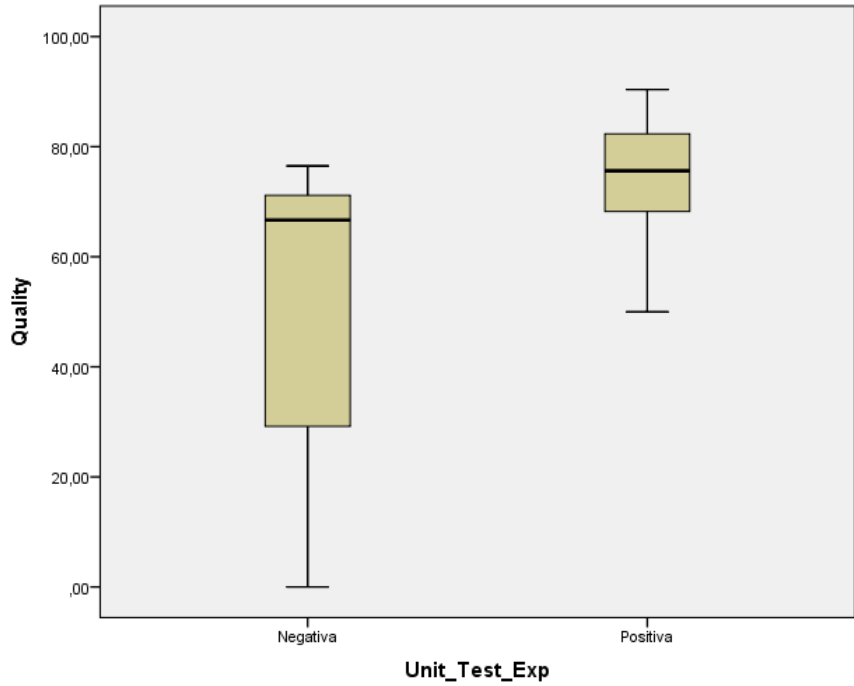
a. Corrección de la significación de Lilliefors

La variable Experiencia en pruebas de unidad no supera el test de normalidad.

2.14.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Unit Test Exp			Estadístico	Error típ.
Quality	Negativa	Media	51,3063	11,37630
		Intervalo de confianza para la media al 95%	Límite inferior	24,4056
			Límite superior	78,2069
		Media recortada al 5%	52,7569	
		Mediana	66,6700	
		Varianza	1035,361	
		Desv. típ.	32,17703	
		Mínimo	,00	
		Máximo	76,50	
		Rango	76,50	
		Amplitud intercuartil	58,79	
		Asimetría	-1,307	,752
		Curtosis	-,173	1,481
	Positiva	Media	74,2981	2,80529
		Intervalo de confianza para la media al 95%	Límite inferior	68,3188
			Límite superior	80,2775
		Media recortada al 5%	74,7551	
		Mediana	75,6050	
		Varianza	125,914	
		Desv. típ.	11,22115	
		Mínimo	50,00	
		Máximo	90,37	
		Rango	40,37	
		Amplitud intercuartil	16,30	
		Asimetría	-,671	,564
		Curtosis	-,002	1,091

2.14.3 DIAGRAMA DE CAJAS: TUS



2.14.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de Unit_Test_Exp.	Prueba U de Mann-Whitney de muestras independientes	,045 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Se puede decir que la variable Experiencia en pruebas de unidad influye sobre la variable respuesta: Calidad.

2.15 CALIDAD – EXPERIENCIA DE PROGRAMACIÓN

2.15.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Programacion_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,305	4	.	,789	4	,084
	Positiva	,124	20	,200 [*]	,974	20	,836

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

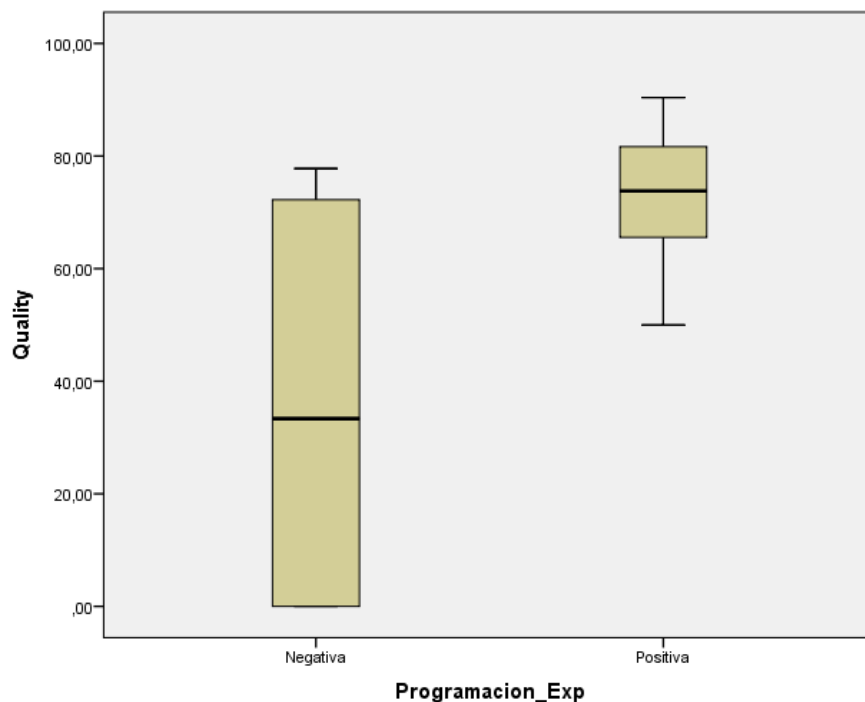
La variable Experiencia de programación supera el test de normalidad.

2.15.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Programacion_Exp				Estadístico	Error típ.
Quality	Negativa	Media		36,1125	20,97253
		Intervalo de confianza para la media al 95%	Límite inferior	-30,6315	
			Límite superior	102,8565	
		Media recortada al 5%		35,8039	
		Mediana		33,3350	
		Varianza		1759,389	
		Desv. típ.		41,94507	
		Mínimo		,00	
		Máximo		77,78	
		Rango		77,78	
		Amplitud intercuartil		75,00	
		Asimetría		,060	1,014
		Curtosis		-5,652	2,619
	Positiva	Media		72,7385	2,40934
		Intervalo de confianza para la media al 95%	Límite inferior	67,6957	
			Límite superior	77,7813	
		Media recortada al 5%		73,0222	
		Mediana		73,7900	
		Varianza		116,098	
		Desv. típ.		10,77489	
		Mínimo		50,00	
		Máximo		90,37	
		Rango		40,37	
		Amplitud intercuartil		16,67	
		Asimetría		-,329	,512
		Curtosis		-,402	,992

2.15.3 DIAGRAMA DE CAJAS



2.15.4 RESULTADOS

Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia
Quality	Se han asumido varianzas iguales	67,570	,000	-3,626	22	,001	-36,62600	10,10223	-57,57674 -15,67526
	No se han asumido varianzas iguales			-1,735	3,080	,179	-36,62600	21,11047	-102,83690 29,58490

2.16 CALIDAD – METODOLOGÍAS DE DESARROLLO

2.16.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	MetDes_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,267	5	,200*	,847	5	,185
	Positiva	,222	19	,015	,709	19	,000

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

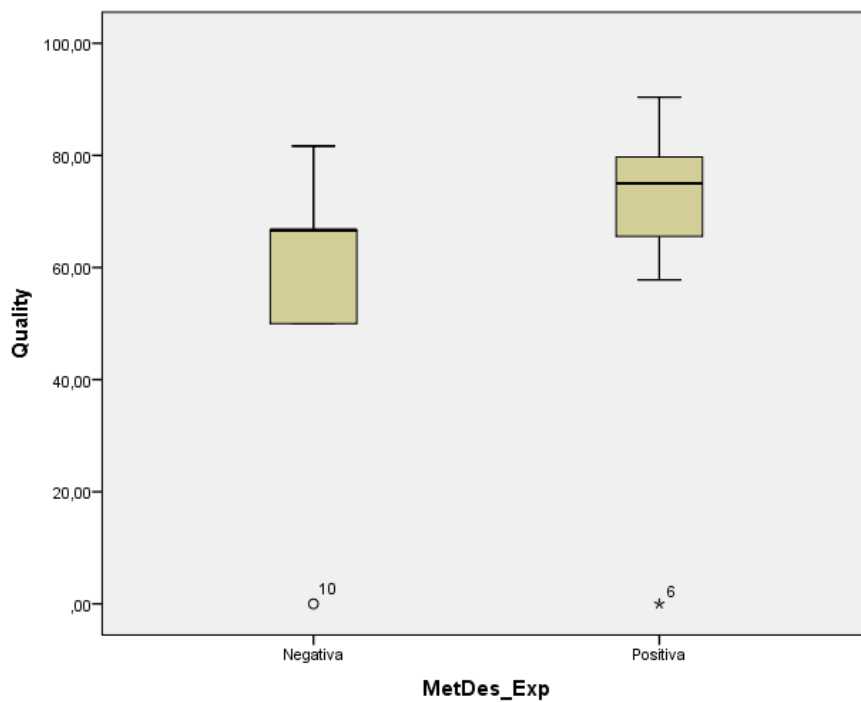
La variable Metodologías de desarrollo no supera el test de normalidad.

2.16.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

MetDes Exp			Estadístico	Error tip.
Quality	Negativa	Media	53,0020	14,16635
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	13,6699 92,3341
		Media recortada al 5%	54,3539	
		Mediana	66,6700	
		Varianza	1003,427	
		Desv. tip.	31,67691	
		Mínimo	,00	
		Máximo	81,67	
		Rango	81,67	
		Amplitud intercuartil	49,17	
		Asimetría	-1,576	,913
		Curtosis	2,723	2,000
	Positiva	Media	70,2216	4,44859
		Intervalo de confianza para la media al 95%	Límite inferior Límite superior	60,8754 79,5677
		Media recortada al 5%	73,0034	
		Mediana	75,0000	
		Varianza	376,010	
		Desv. tip.	19,39097	
		Mínimo	,00	
		Máximo	90,37	
		Rango	90,37	
		Amplitud intercuartil	17,23	
		Asimetría	-2,802	,524
		Curtosis	10,004	1,014

2.16.3 DIAGRAMA DE CAJAS



2.16.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de Quality es la misma entre las categorías de MetDes_Exp.	Prueba U de Mann-Whitney de muestras independientes	,160 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

2.17 TUS – EXPERIENCIA ACADÉMICA

2.17.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Academic_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,298	4	.	,849	4	,224
	Positiva	,219	20	,013	,812	20	,001

a. Corrección de la significación de Lilliefors

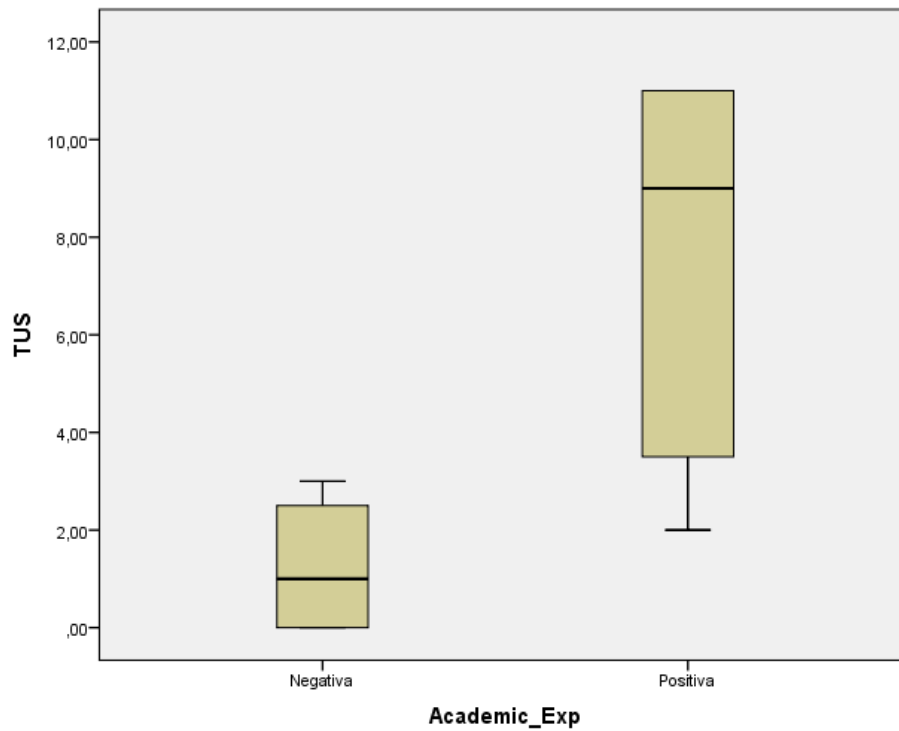
La variable Experiencia académica no supera el test de normalidad.

2.17.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Academic Exp				Estadístico	Error típ.
TUS	Negativa	Media		1,2500	,75000
		Intervalo de confianza para la media al 95%	Límite inferior	-1,1368	
			Límite superior	3,6368	
		Media recortada al 5%		1,2222	
		Mediana		1,0000	
		Varianza		2,250	
		Desv. típ.		1,50000	
		Mínimo		,00	
		Máximo		3,00	
		Rango		3,00	
		Amplitud intercuartil		2,75	
		Asimetría		,370	1,014
		Curtosis		-3,901	2,619
	Positiva	Media		7,4000	,81886
		Intervalo de confianza para la media al 95%	Límite inferior	5,6861	
			Límite superior	9,1139	
		Media recortada al 5%		7,5000	
		Mediana		9,0000	
		Varianza		13,411	
		Desv. típ.		3,66204	
		Mínimo		2,00	
		Máximo		11,00	
		Rango		9,00	
		Amplitud intercuartil		7,75	
		Asimetría		-,456	,512
		Curtosis		-1,639	,992

2.17.3 DIAGRAMA DE CAJAS: TUS



2.17.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Academic_Exp.	Prueba U de Mann-Whitney de muestras independientes	,003 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Según la gráfica la variable Experiencia académica influye sobre la variable respuesta: TUS.

2.18 TUS – EXPERIENCIA DE JAVA

2.18.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Progr_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,329	4	.	,895	4	,406
	Positiva	,228	20	,008	,830	20	,002

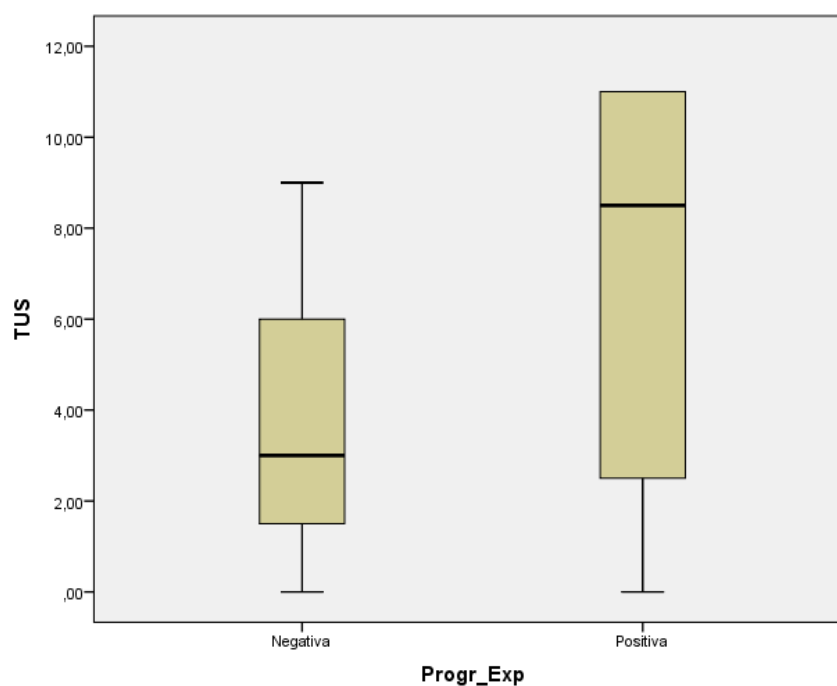
a. Corrección de la significación de Lilliefors

La variable Experiencia de programación no supera el test de normalidad.

2.18.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				Estadístico	Error típ.
Progr. Exp					
TUS	Negativa	Media		3,7500	1,88746
		Intervalo de confianza para la media al 95%	Límite inferior	-2,2567	
			Límite superior	9,7567	
		Media recortada al 5%		3,6667	
		Mediana		3,0000	
		Varianza		14,250	
		Desv. típ.		3,77492	
		Mínimo		,00	
		Máximo		9,00	
		Rango		9,00	
		Amplitud intercuartil		6,75	
		Asimetría		1,129	1,014
		Curtosis		2,227	2,619
	Positiva	Media		6,9000	,90583
		Intervalo de confianza para la media al 95%	Límite inferior	5,0041	
			Límite superior	8,7959	
		Media recortada al 5%		7,0556	
		Mediana		8,5000	
		Varianza		16,411	
		Desv. típ.		4,05099	
		Mínimo		,00	
		Máximo		11,00	
		Rango		11,00	
		Amplitud intercuartil		8,75	
		Asimetría		-,347	,512
		Curtosis		-1,678	,992

2.18.3 DIAGRAMA DE CAJAS: TUS



2.18.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Progr_Exp.	Prueba U de Mann-Whitney de muestras independientes	,183 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La diferencia de medias propone que la variable experiencia de programación influye en la variable respuesta: TUS.

2.19 TUS – EXPERIENCIA EN TDD

2.19.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	TDD_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,180	12	,200*	,878	12	,083
	Positiva	,237	12	,061	,787	12	,007

*. Este es un límite inferior de la significación verdadera.

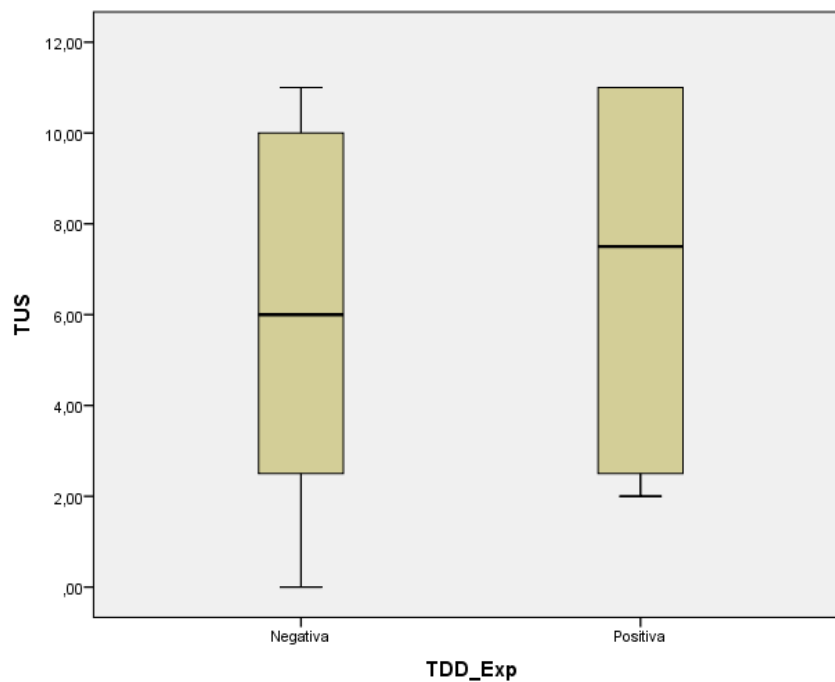
a. Corrección de la significación de Lilliefors

La variable Experiencia en TDD no supera el test de normalidad.

2.19.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
TDD_Exp			Estadístico	Error típ.	
TUS	Negativa	Media	6,0000	1,23091	
		Intervalo de confianza para la media al 95%	Límite inferior	3,2908	
			Límite superior	8,7092	
		Media recortada al 5%	6,0556		
		Mediana	6,0000		
		Varianza	18,182		
		Desv. típ.	4,26401		
		Mínimo	,00		
		Máximo	11,00		
		Rango	11,00		
		Amplitud intercuartil	7,75		
		Asimetría	-,177	,637	
		Curtosis	-1,736	1,232	
	Positiva	Media	6,7500	1,18145	
		Intervalo de confianza para la media al 95%	Límite inferior	4,1496	
			Límite superior	9,3504	
		Media recortada al 5%	6,7778		
		Mediana	7,5000		
		Varianza	16,750		
		Desv. típ.	4,09268		
		Mínimo	2,00		
		Máximo	11,00		
		Rango	9,00		
		Amplitud intercuartil	8,75		
		Asimetría	-,119	,637	
		Curtosis	-2,143	1,232	

2.19.3 DIAGRAMA DE CAJAS: TUS



2.19.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de TDD_Exp.	Prueba U de Mann-Whitney de muestras independientes	,630 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La diferencia de medias no resulta significativa como para afirmar que la variable Experiencia en TDD influye sobre la variable respuesta: TUS.

2.20 TUS – EXPERIENCIA EN JUNIT

2.20.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Junit_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,261	9	,077	,845	9	,065
	Positiva	,226	15	,037	,844	15	,014

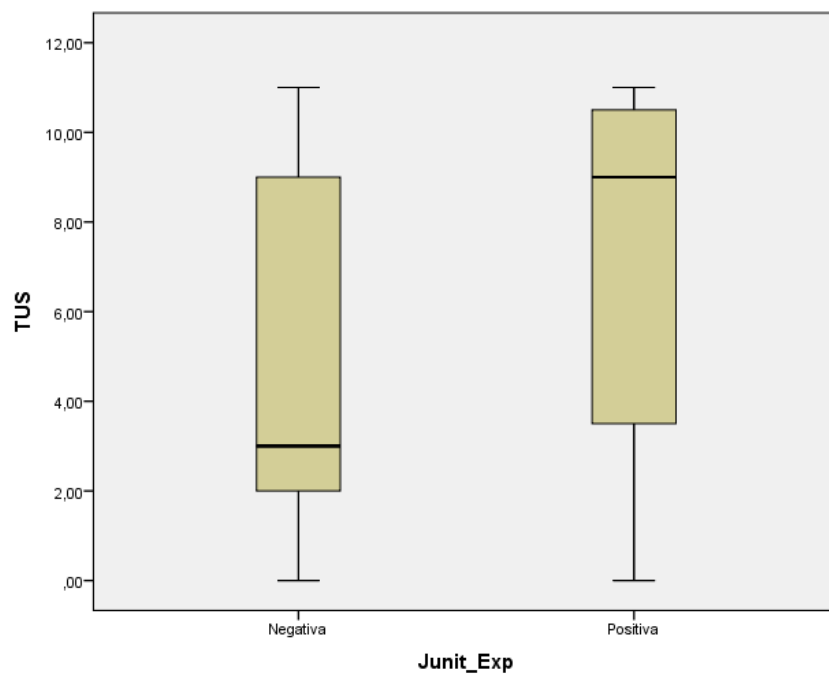
a. Corrección de la significación de Lilliefors

La variable experiencia en JUnit no supera el test de normalidad.

2.20.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				Estadístico	Error típ.
Junit_Exp					
TUS	Negativa	Media		5,0000	1,39443
		Intervalo de confianza para la media al 95%	Límite inferior	1,7844	
			Límite superior	8,2156	
		Media recortada al 5%		4,9444	
		Mediana		3,0000	
		Varianza		17,500	
		Desv. típ.		4,18330	
		Mínimo		,00	
		Máximo		11,00	
		Rango		11,00	
		Amplitud intercuartil		8,00	
		Asimetría		,659	,717
		Curtosis		-1,363	1,400
	Positiva	Media		7,2000	1,02446
		Intervalo de confianza para la media al 95%	Límite inferior	5,0027	
			Límite superior	9,3973	
		Media recortada al 5%		7,3889	
		Mediana		9,0000	
		Varianza		15,743	
		Desv. típ.		3,96773	
		Mínimo		,00	
		Máximo		11,00	
		Rango		11,00	
		Amplitud intercuartil		8,00	
		Asimetría		-,622	,580
		Curtosis		-1,287	1,121

2.20.3 DIAGRAMA DE CAJAS: TUS



2.20.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Junit_Exp.	Prueba U de Mann-Whitney de muestras independientes	,290 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La variable Experiencia en JUnit influye sobre la variable respuesta: TUS

2.21 TUS – EXPERIENCIA IDE

2.21.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	IDE_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,271	4	.	,848	4	,220
	Positiva	,217	20	,015	,841	20	,004

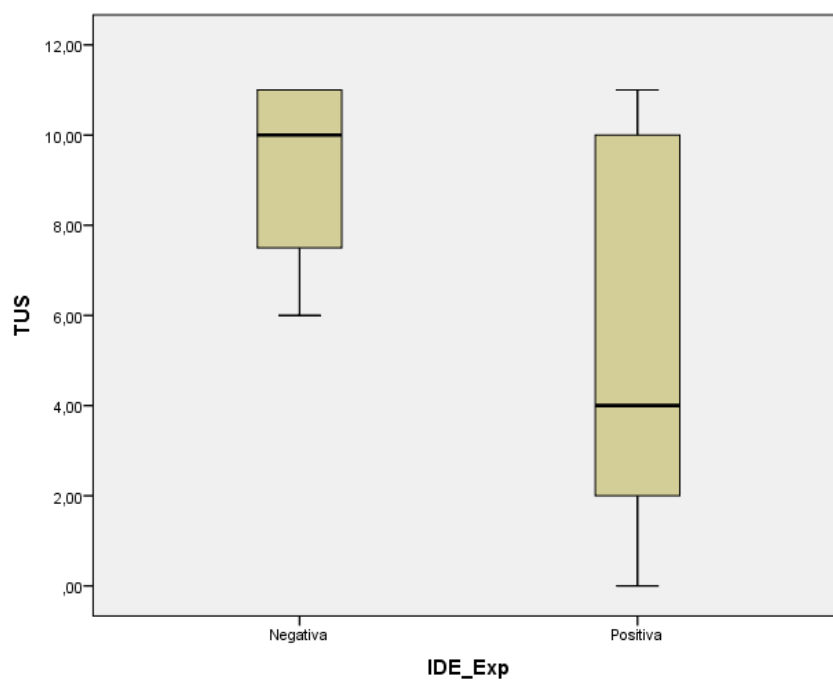
a. Corrección de la significación de Lilliefors

La variable experiencia en el uso de IDEs no supera el test de normalidad.

2.21.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
IDE_Exp			Estadístico	Error típ.	
TUS	Negativa	Media	9,2500	1,18145	
		Intervalo de confianza para la media al 95%	Límite inferior	5,4901	
			Límite superior	13,0099	
		Media recortada al 5%	9,3333		
		Mediana	10,0000		
		Varianza	5,583		
		Desv. típ.	2,36291		
		Mínimo	6,00		
		Máximo	11,00		
		Rango	5,00		
		Amplitud intercuartil	4,25		
		Asimetría	-1,194	1,014	
		Curtosis	,436	2,619	
	Positiva	Media	5,8000	,93358	
		Intervalo de confianza para la media al 95%	Límite inferior	3,8460	
			Límite superior	7,7540	
		Media recortada al 5%	5,8333		
		Mediana	4,0000		
		Varianza	17,432		
		Desv. típ.	4,17511		
		Mínimo	,00		
		Máximo	11,00		
		Rango	11,00		
		Amplitud intercuartil	8,00		
		Asimetría	,113	,512	
		Curtosis	-1,806	,992	

2.21.3 DIAGRAMA DE CAJAS: TUS



2.21.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de IDE_Exp.	Prueba U de Mann-Whitney de muestras independientes	,135 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Nuevamente, la variable Experiencia en el uso de IDEs nos presenta un dato curioso, resulta influyente sobre la variable respuesta TUS, pero los usuarios con menos experiencia presenta mayor TUS.

2.22 TUS – EXPERIENCIA PRUEBAS DE UNIDAD

2.22.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Unit_Test_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,250	8	,150	,830	8	,059
	Positiva	,218	16	,041	,820	16	,005

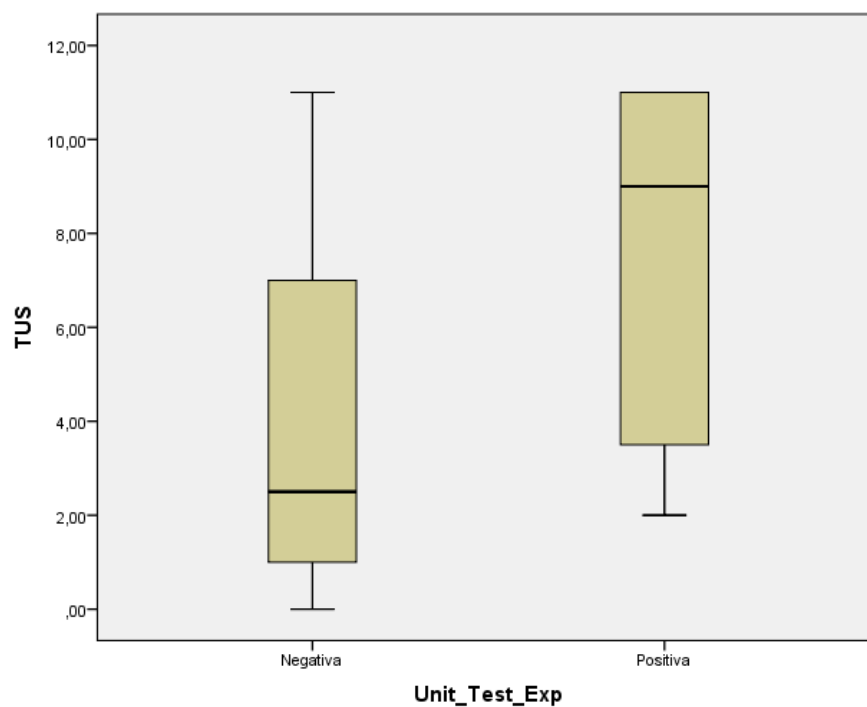
a. Corrección de la significación de Lilliefors

La variable Experiencia en pruebas de unidad no supera el test de normalidad.

2.22.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Unit Test Exp			Estadístico	Error típ.
TUS	Negativa	Media	4,0000	1,50000
		Intervalo de confianza para la media al 95%	Límite inferior	,4531
			Límite superior	7,5469
		Media recortada al 5%	3,8333	
		Mediana	2,5000	
		Varianza	18,000	
		Desv. típ.	4,24264	
		Mínimo	,00	
		Máximo	11,00	
		Rango	11,00	
		Amplitud intercuartil	8,00	
		Asimetría	1,033	,752
		Curtosis	-,411	1,481
	Positiva	Media	7,5625	,89893
		Intervalo de confianza para la media al 95%	Límite inferior	5,6465
			Límite superior	9,4785
		Media recortada al 5%	7,6806	
		Mediana	9,0000	
		Varianza	12,929	
		Desv. típ.	3,59572	
		Mínimo	2,00	
		Máximo	11,00	
		Rango	9,00	
		Amplitud intercuartil	7,75	
		Asimetría	-,566	,564
		Curtosis	-1,488	1,091

2.22.3 DIAGRAMA DE CAJAS: TUS



2.22.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Unit_Test_Exp.	Prueba U de Mann-Whitney de muestras independientes	,061 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Dada la diferencia de medias, hay clara evidencia de que la variable Experiencia en pruebas de unidad influye sobre la variable respuesta: TUS.

2.23 TUS – EXPERIENCIA DE PROGRAMACIÓN

2.23.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Programacion_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,298	4	.	,849	4	,224
	Positiva	,219	20	,013	,812	20	,001

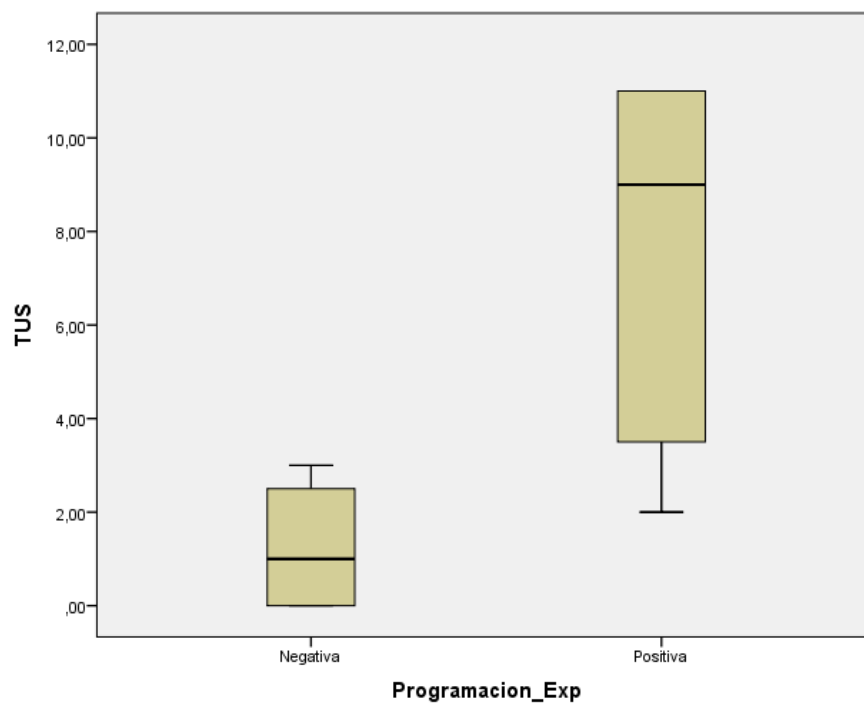
a. Corrección de la significación de Lilliefors

La variable experiencia de programación no supera el test de normalidad.

2.23.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Programacion_Exp				Estadístico	Error típ.
TUS	Negativa	Media		1,2500	,75000
		Intervalo de confianza para la media al 95%	Límite inferior	-1,1368	
			Límite superior	3,6368	
		Media recortada al 5%		1,2222	
		Mediana		1,0000	
		Varianza		2,250	
		Desv. típ.		1,50000	
		Mínimo		,00	
		Máximo		3,00	
		Rango		3,00	
		Amplitud intercuartil		2,75	
		Asimetría		,370	1,014
		Curtosis		-3,901	2,619
	Positiva	Media		7,4000	,81886
		Intervalo de confianza para la media al 95%	Límite inferior	5,6861	
			Límite superior	9,1139	
		Media recortada al 5%		7,5000	
		Mediana		9,0000	
		Varianza		13,411	
		Desv. típ.		3,66204	
		Mínimo		2,00	
		Máximo		11,00	
		Rango		9,00	
		Amplitud intercuartil		7,75	
		Asimetría		-,456	,512
		Curtosis		-1,639	,992

2.23.3 DIAGRAMA DE CAJAS



2.23.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Programacion_Exp.	Prueba U de Mann-Whitney de muestras independientes	,003 ¹	Rechazar la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

2.24 TUS – METODOLOGÍAS DE DESARROLLO

2.24.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	MetDes_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,356	5	,038	,779	5	,054
	Positiva	,215	19	,021	,853	19	,008

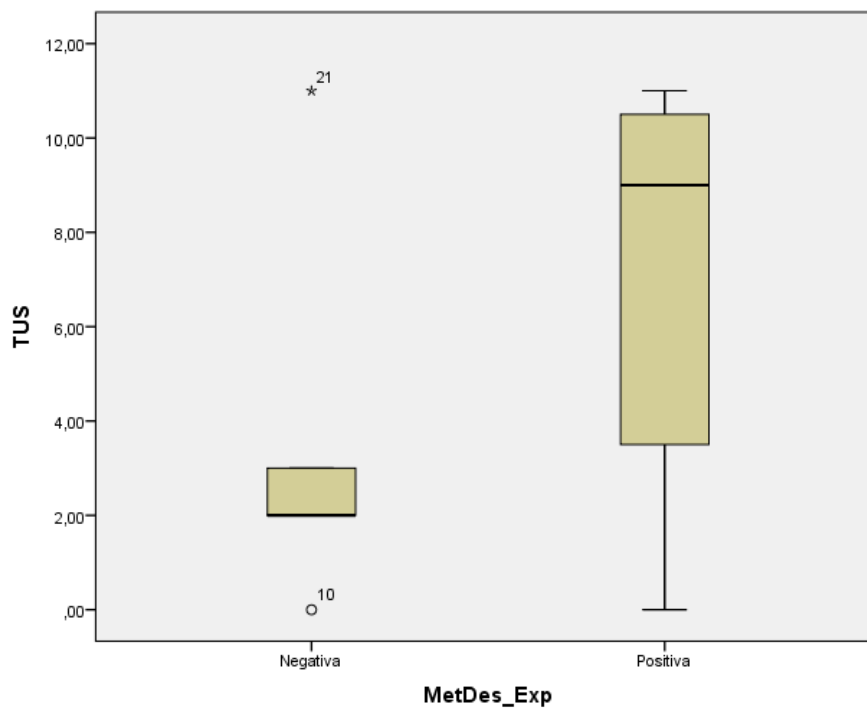
a. Corrección de la significación de Lilliefors

La variable Metodologías de desarrollo no supera el test de normalidad.

2.24.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				Estadístico	Error típ.
TUS	MetDes	Exp			
	Negativa	Media		3,6000	1,91311
		Intervalo de confianza para la media al 95%	Límite inferior	-1,7117	
			Límite superior	8,9117	
		Media recortada al 5%		3,3889	
		Mediana		2,0000	
		Varianza		18,300	
		Desv. típ.		4,27785	
		Mínimo		,00	
		Máximo		11,00	
		Rango		11,00	
		Amplitud intercuartil		6,00	
		Asimetría		1,864	,913
		Curtosis		3,869	2,000
	Positiva	Media		7,1053	,88157
		Intervalo de confianza para la media al 95%	Límite inferior	5,2532	
			Límite superior	8,9574	
		Media recortada al 5%		7,2836	
		Mediana		9,0000	
		Varianza		14,766	
		Desv. típ.		3,84267	
		Mínimo		,00	
		Máximo		11,00	
		Rango		11,00	
		Amplitud intercuartil		8,00	
		Asimetría		-,484	,524
		Curtosis		-1,420	1,014

2.24.3 DIAGRAMA DE CAJAS



2.24.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de MetDes_Exp.	Prueba U de Mann-Whitney de muestras independientes	,120 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

3 ANÁLISIS DICOTÓMICO MUSICPHONE

3.1 PRODUCTIVIDAD – EXPERIENCIA ACADÉMICA

3.1.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Academic_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,321	3	.	,881	3	,328
	Positiva	,115	20	,200*	,977	20	,890

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

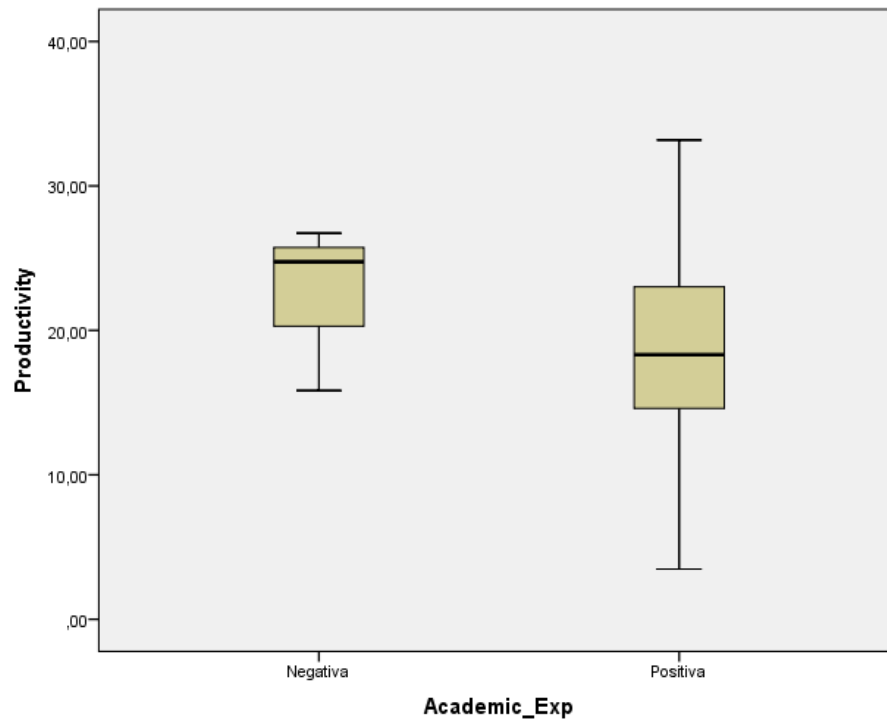
Como se puede observar la variable experiencia académica supera el test de normalidad.

3.1.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Academic Exp				Estadístico	Error típ.
Productivity	Negativa	Media		22,4400	3,34913
		Intervalo de confianza para la media al 95%	Límite inferior	8,0298	
			Límite superior	36,8502	
		Media recortada al 5%		.	
		Mediana		24,7500	
		Varianza		33,650	
		Desv. típ.		5,80087	
		Mínimo		15,84	
		Máximo		26,73	
		Rango		10,89	
		Amplitud intercuartil		.	
		Asimetría		-1,508	1,225
		Curtosis		.	.
	Positiva	Media		18,9860	1,65493
		Intervalo de confianza para la media al 95%	Límite inferior	15,5222	
			Límite superior	22,4498	
		Media recortada al 5%		19,0600	
		Mediana		18,3200	
		Varianza		54,776	
		Desv. típ.		7,40106	
		Mínimo		3,47	
		Máximo		33,17	
		Rango		29,70	
		Amplitud intercuartil		8,66	
		Asimetría		,078	,512
		Curtosis		-,035	,992

3.1.3 DIAGRAMA DE CAJAS: TUS



3.1.4 RESULTADOS

Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia
Productivity	Se han asumido varianzas iguales	,371	,549	,768	21	,451	3,45400	4,49735	-5,89874 12,80674
	No se han asumido varianzas iguales			,925	3,077	,422	3,45400	3,73570	-8,26894 15,17694

La variable experiencia académica es influyente sobre la variable respuesta: Productividad.

3.2 PRODUCTIVIDAD – EXPERIENCIA JAVA

3.2.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Progr_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,178	3	.	,999	3	,952
	Positiva	,111	20	,200 [*]	,974	20	,843

*. Este es un límite inferior de la significación verdadera.

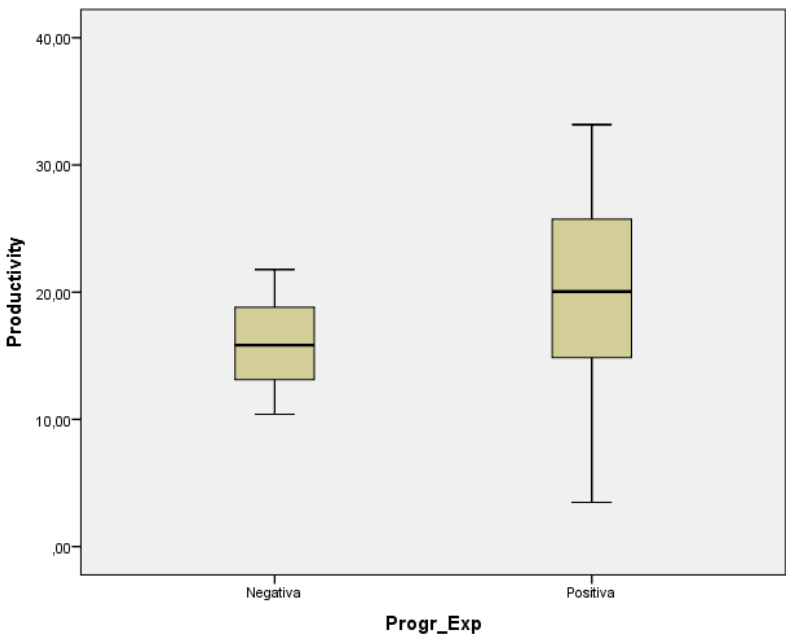
a. Corrección de la significación de Lilliefors

Como se puede observar la variable Experiencia de programación supera el test de normalidad.

3.2.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Progr. Exp				Estadístico	Error típ.
Productivity	Negativa	Media		16,0067	3,28618
		Intervalo de confianza para la media al 95%	Límite inferior	1,8674	
			Límite superior	30,1460	
		Media recortada al 5%			
		Mediana		15,8400	
		Varianza		32,397	
		Desv. típ.		5,69183	
		Mínimo		10,40	
		Máximo		21,78	
		Rango		11,38	
		Amplitud intercuartil			
		Asimetría		,132	1,225
		Curtosis			
	Positiva	Media		19,9510	1,64939
		Intervalo de confianza para la media al 95%	Límite inferior	16,4988	
			Límite superior	23,4032	
		Media recortada al 5%		20,1322	
		Mediana		20,0500	
		Varianza		54,410	
		Desv. típ.		7,37629	
		Mínimo		3,47	
		Máximo		33,17	
		Rango		29,70	
		Amplitud intercuartil		11,39	
		Asimetría		-,175	,512
		Curtosis		-,035	,992

3.2.3 DIAGRAMA DE CAJAS: TUS



3.2.4 RESULTADOS

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia	
Productivity	Se han asumido varianzas iguales	,808	,379	-,881	21	,388	-3,94433	4,47810	-13,25705	5,36838
	No se han asumido varianzas iguales			-1,073	3,114	,359	-3,94433	3,67688	-15,40755	7,51888

Dada la diferencia de medias la variable Experiencia de programación tiene influencia sobre la variable respuesta: Productividad.

3.3 PRODUCTIVIDAD – EXPERIENCIA EN TDD

3.3.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	TDD_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,195	11	,200 [*]	,941	11	,535
	Positiva	,143	12	,200 [*]	,983	12	,992

*. Este es un límite inferior de la significación verdadera.

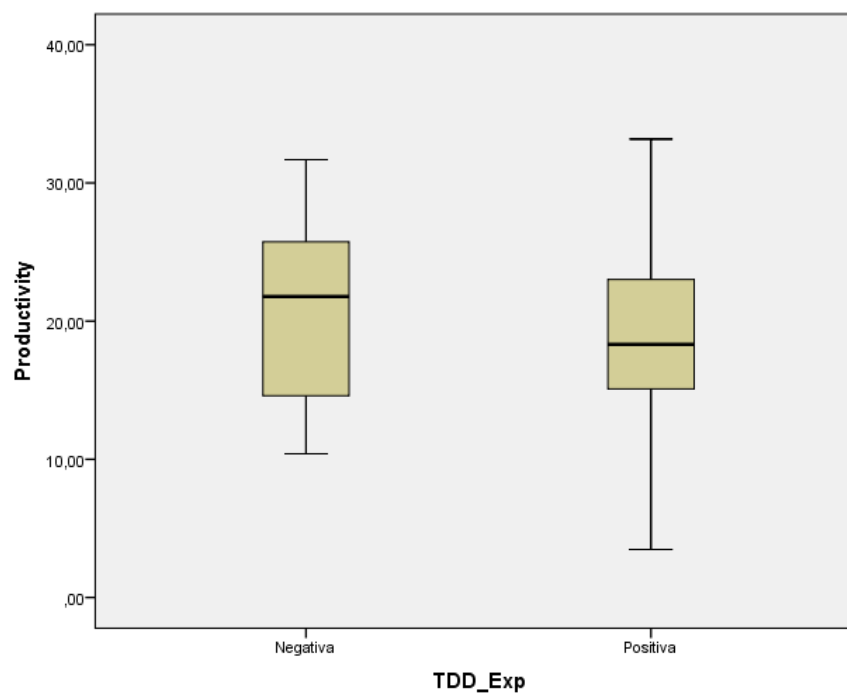
a. Corrección de la significación de Lilliefors

La variable Experiencia en TDD supera el test de normalidad.

3.3.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
TDD Exp				Estadístico	Error típ.
Productivity	Negativa	Media		20,2518	2,06485
		Intervalo de confianza para la media al 95%	Límite inferior	15,6511	
			Límite superior	24,8526	
		Media recortada al 5%		20,1642	
		Mediana		21,7800	
		Varianza		46,899	
		Desv. típ.		6,84832	
		Mínimo		10,40	
		Máximo		31,68	
		Rango		21,28	
		Amplitud intercuartil		12,37	
		Asimetría		,156	,661
		Curtosis		-1,240	1,279
	Positiva	Media		18,6892	2,22954
		Intervalo de confianza para la media al 95%	Límite inferior	13,7820	
			Límite superior	23,5964	
		Media recortada al 5%		18,7302	
		Mediana		18,3200	
		Varianza		59,650	
		Desv. típ.		7,72336	
		Mínimo		3,47	
		Máximo		33,17	
		Rango		29,70	
		Amplitud intercuartil		8,17	
		Asimetría		-,109	,637
		Curtosis		,738	1,232

3.3.3 DIAGRAMA DE CAJAS: TUS



3.3.4 RESULTADOS

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia	
Productivity	Se han asumido varianzas iguales	,012	,914	,511	21	,614	1,56265	3,05543	-4,79146	7,91676
	No se han asumido varianzas iguales			,514	20,982	,612	1,56265	3,03882	-4,75725	7,88255

La variable Experiencia en TDD no presenta una diferencia de medias significativa como para asegurar que influye sobre la variable respuesta Productividad. Habría que realizar más pruebas.

3.4 PRODUCTIVIDAD – EXPERIENCIA EN JUNIT

3.4.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Junit_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,161	8	,200*	,968	8	,884
	Positiva	,134	15	,200*	,976	15	,930

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

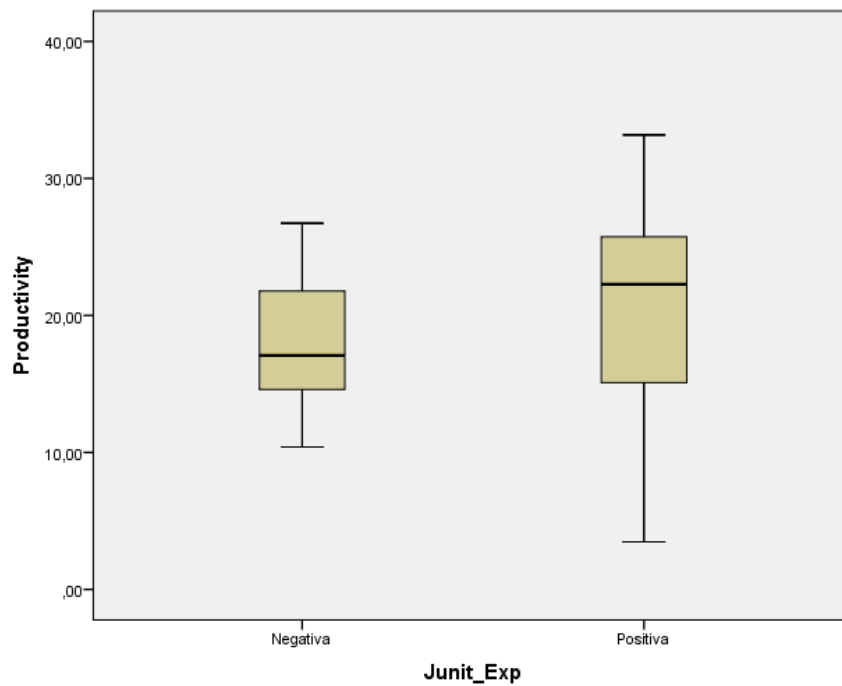
La variable experiencia en JUnit supera el test de normalidad.

3.4.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Junit_Exp			Estadístico	Error típ.
Productivity	Negativa	Media	18,0075	1,84434
		Intervalo de confianza para la media al 95%	Límite inferior	13,6463
			Límite superior	22,3687
		Media recortada al 5%	17,9456	
		Mediana	17,0800	
		Varianza	27,213	
		Desv. típ.	5,21659	
		Mínimo	10,40	
		Máximo	26,73	
		Rango	16,33	
		Amplitud intercuartil	7,30	
		Asimetría	,323	,752
		Curtosis	-,344	1,481
	Positiva	Media	20,1987	2,09719
		Intervalo de confianza para la media al 95%	Límite inferior	15,7006
			Límite superior	24,6967
		Media recortada al 5%	20,4074	
		Mediana	22,2800	
		Varianza	65,973	
		Desv. típ.	8,12238	
		Mínimo	3,47	
		Máximo	33,17	
		Rango	29,70	
		Amplitud intercuartil	11,88	
		Asimetría	-,291	,580
		Curtosis	-,200	1,121

3.4.3 DIAGRAMA DE CAJAS: TUS



3.4.4 RESULTADOS

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia	
Productivity	Se han asumido varianzas iguales	2,159	,157	-,687	21	,500	-2,19117	3,18881	-8,82265	4,44032
	No se han asumido varianzas iguales			-,785	20,047	,442	-2,19117	2,79281	-8,01599	3,63366

Dada la diferencia de medias, la variable Experiencia en JUnit influye sobre la variable respuesta: Productividad.

3.5 PRODUCTIVIDAD – EXPERIENCIA IDE

3.5.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	IDE_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,134	4	.	,998	4	,995
	Positiva	,157	19	,200 [*]	,956	19	,493

*. Este es un límite inferior de la significación verdadera.

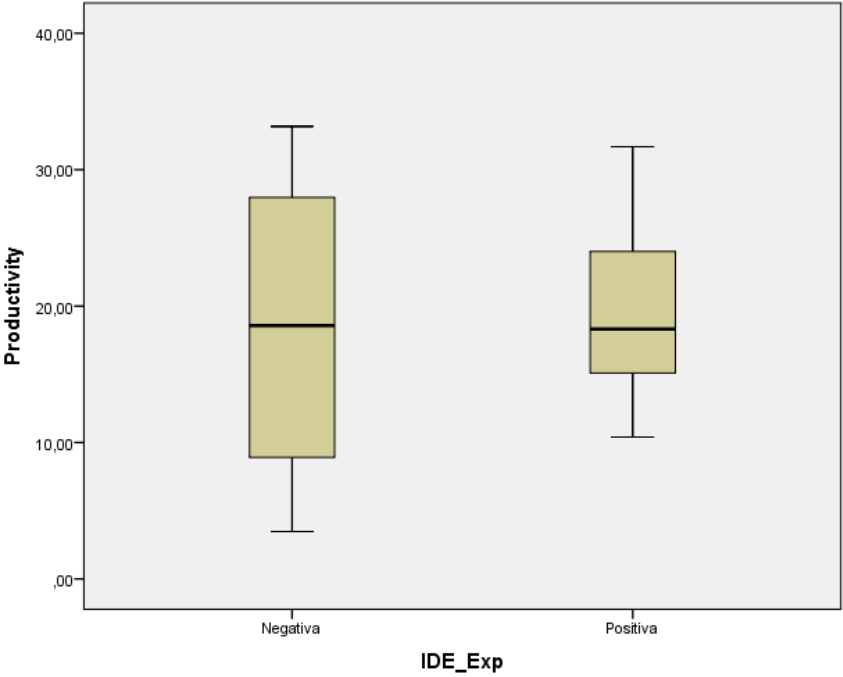
a. Corrección de la significación de Lilliefors

La variable Experiencia en el uso de IDEs supera el test de normalidad.

3.5.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
IDE_Exp				Estadístico	Error típ.
Productivity	Negativa	Media		18,4425	6,30125
		Intervalo de confianza para la media al 95%	Límite inferior	-1,6109	
			Límite superior	38,4959	
		Media recortada al 5%		18,4561	
		Mediana		18,5650	
		Varianza		158,823	
		Desv. típ.		12,60250	
		Mínimo		3,47	
		Máximo		33,17	
		Rango		29,70	
		Amplitud intercuartil		24,38	
		Asimetría		-,050	1,014
		Curtosis		-,559	2,619
	Positiva	Media		19,6458	1,38698
		Intervalo de confianza para la media al 95%	Límite inferior	16,7319	
			Límite superior	22,5597	
		Media recortada al 5%		19,4909	
		Mediana		18,3200	
		Varianza		36,550	
		Desv. típ.		6,04570	
		Mínimo		10,40	
		Máximo		31,68	
		Rango		21,28	
		Amplitud intercuartil		9,90	
		Asimetría		,221	,524
		Curtosis		-,816	1,014

3.5.3 DIAGRAMA DE CAJAS: TUS



3.5.4 RESULTADOS

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia	
Productivity	Se han asumido varianzas iguales	4,973	,037	-,298	21	,769	-1,20329	4,04321	-9,61160	7,20502
	No se han asumido varianzas iguales			-,186	3,296	,863	-1,20329	6,45209	-20,73060	18,32402

La diferencia de medias no es significativa, por lo tanto no se puede asegurar que la variable Experiencia en el uso de IDEs influye sobre la variable respuesta: Productividad.

3.6 PRODUCTIVIDAD – EXPERIENCIA PRUEBAS DE UNIDAD

3.6.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
Unit_Test_Exp		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,308	7	,043	,825	7	,072
	Positiva	,109	16	,200*	,982	16	,979

*. Este es un límite inferior de la significación verdadera.

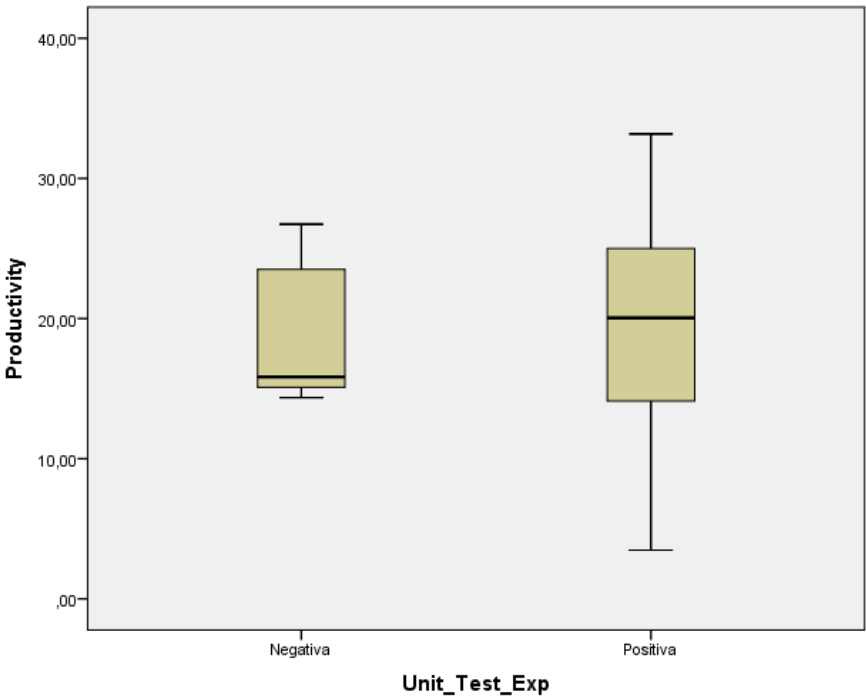
a. Corrección de la significación de Lilliefors

La variable Experiencia en pruebas de unidad supera el test de normalidad.

3.6.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Unit Test Exp			Estadístico	Error tip.
Productivity	Negativa	Media	19,1657	1,98471
		Intervalo de confianza para la media al 95%	Límite inferior	14,3093
			Límite superior	24,0221
		Media recortada al 5%	19,0125	
		Mediana	15,8400	
		Varianza	27,574	
		Desv. tip.	5,25106	
		Mínimo	14,36	
		Máximo	26,73	
		Rango	12,37	
		Amplitud intercuartil	9,90	
		Asimetría	,570	,794
		Curtosis	-1,990	1,587
	Positiva	Media	19,5550	2,01348
		Intervalo de confianza para la media al 95%	Límite inferior	15,2634
			Límite superior	23,8466
		Media recortada al 5%	19,6922	
		Mediana	20,0500	
		Varianza	64,865	
		Desv. tip.	8,05391	
		Mínimo	3,47	
		Máximo	33,17	
		Rango	29,70	
		Amplitud intercuartil	12,13	
		Asimetría	-,137	,564
		Curtosis	-,281	1,091

3.6.3 DIAGRAMA DE CAJAS: TUS



Resultados

Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia
									Inferior Superior
Productivity	Se han asumido varianzas iguales	1,005	,328	-,117	21	,908	-,38929	3,33655	-7,32801 6,54944
	No se han asumido varianzas iguales			-,138	17,353	,892	-,38929	2,82722	-6,34496 5,56639

Dado que la diferencia de medias no es significativa, no hay indicios para determinar que la variable Experiencia en pruebas de unidad influye sobre la variable respuesta: Productividad.

3.7 PRODUCTIVIDAD – EXPERIENCIA DE PROGRAMACIÓN

3.7.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad ^a							
	Programacion_Exp	Kolmogorov-Smirnov ^b			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,281	4	.	,842	4	,202
	Positiva	,107	19	,200 [*]	,981	19	,952

*. Este es un límite inferior de la significación verdadera.

a. No hay ningún caso válido para Productivity cuando Programacion_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

b. Corrección de la significación de Lilliefors

La variable Experiencia de programación supera el test de normalidad.

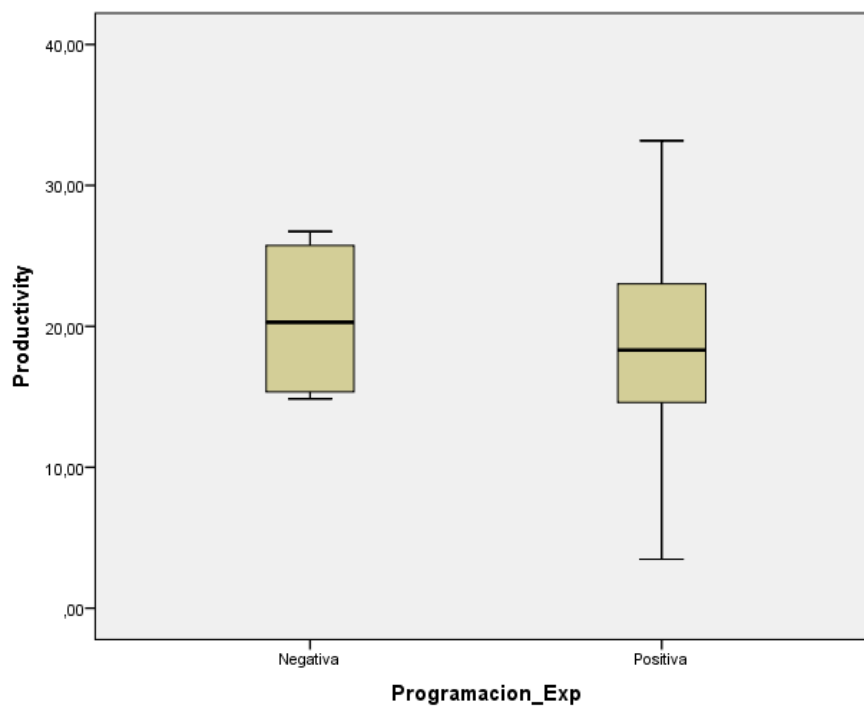
3.7.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos^a

Programacion_Exp				Estadístico	Error típ.
Productivity	Negativa	Media		20,5425	3,03461
		Intervalo de confianza para la media al 95%	Límite inferior	10,8850	
			Límite superior	30,2000	
		Media recortada al 5%		20,5150	
		Mediana		20,2950	
		Varianza		36,835	
		Desv. típ.		6,06922	
		Mínimo		14,85	
		Máximo		26,73	
		Rango		11,88	
		Amplitud intercuartil		11,14	
		Asimetría		,068	1,014
		Curtosis		-5,348	2,619
	Positiva	Media		19,2037	1,72929
		Intervalo de confianza para la media al 95%	Límite inferior	15,5706	
			Límite superior	22,8368	
		Media recortada al 5%		19,3019	
		Mediana		18,3200	
		Varianza		56,818	
		Desv. típ.		7,53780	
		Mínimo		3,47	
		Máximo		33,17	
		Rango		29,70	
		Amplitud intercuartil		8,91	
		Asimetría		-,008	,524
		Curtosis		-,087	1,014

a. No hay ningún caso válido para Productivity cuando Programacion_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

3.7.3 DIAGRAMA DE CAJAS



3.7.4 RESULTADOS

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia	
Productivity	Se han asumido varianzas iguales	,147	,705	,331	21	,744	1,33882	4,04118	-7,06528	9,74291
	No se han asumido varianzas iguales			,383	5,174	,717	1,33882	3,49275	-7,54960	10,22723

3.8 PRODUCTIVIDAD – METODOLOGÍAS DE DESARROLLO

3.8.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad ^a							
	MetDes_Exp	Kolmogorov-Smirnov ^b			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Productivity	Negativa	,186	4	.	,989	4	,952
	Positiva	,157	19	,200*	,973	19	,840

*. Este es un límite inferior de la significación verdadera.

a. No hay ningún caso válido para Productivity cuando MetDes_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

b. Corrección de la significación de Lilliefors

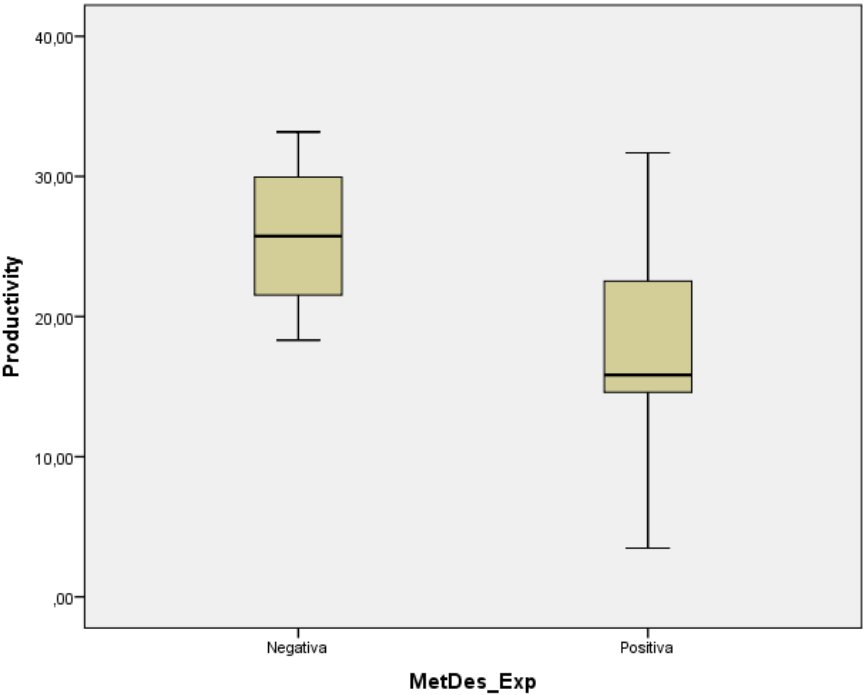
La variable Metodologías de desarrollo supera el test de normalidad.

3.8.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos ^a					
MetDes_Exp				Estadístico	Error típ.
Productivity	Negativa	Media		25,7425	3,05807
		Intervalo de confianza para la media al 95%	Límite inferior	16,0104	
			Límite superior	35,4746	
		Media recortada al 5%		25,7422	
		Mediana		25,7400	
		Varianza		37,407	
		Desv. típ.		6,11614	
		Mínimo		18,32	
		Máximo		33,17	
		Rango		14,85	
		Amplitud intercuartil		11,63	
		Asimetría		,002	1,014
		Curtosis		,985	2,619
	Positiva	Media		18,1089	1,56197
		Intervalo de confianza para la media al 95%	Límite inferior	14,8274	
			Límite superior	21,3905	
		Media recortada al 5%		18,1683	
		Mediana		15,8400	
		Varianza		46,355	
		Desv. típ.		6,80846	
		Mínimo		3,47	
		Máximo		31,68	
		Rango		28,21	
		Amplitud intercuartil		8,41	
		Asimetría		-,027	,524
		Curtosis		,069	1,014

a. No hay ningún caso válido para Productivity cuando MetDes_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

3.8.3 DIAGRAMA DE CAJAS



3.8.4 RESULTADOS

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia	
Productivity	Se han asumido varianzas iguales	,395	,536	2,067	21	,051	7,63355	3,69346	-,04743	15,31453
	No se han asumido varianzas iguales			2,223	4,716	,080	7,63355	3,43388	-1,35582	16,62293

3.9 CALIDAD – EXPERIENCIA ACADÉMICA

3.9.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Experience	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,185	3	.	,998	3	,922
	Positiva	,114	20	,200*	,984	20	,975

*. Este es un límite inferior de la significación verdadera.

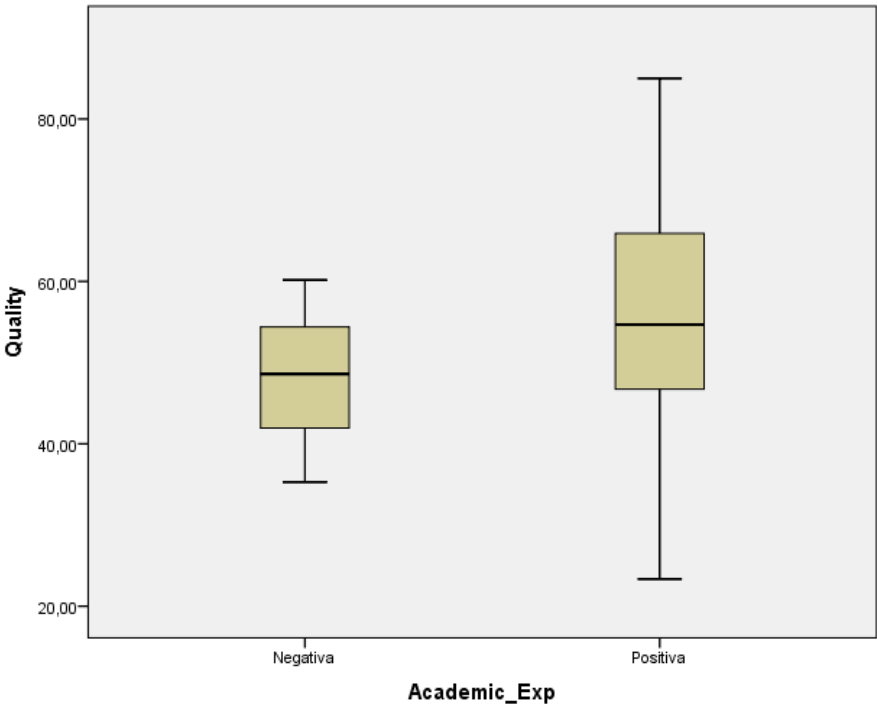
a. Corrección de la significación de Lilliefors

La variable Experiencia académica supera el test de normalidad.

3.9.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Academic Exp			Estadístico	Error típ.
Quality	Negativa	Media	48,0233	7,19976
		Intervalo de confianza para la media al 95%	Límite inferior	17,0453
			Límite superior	79,0014
		Media recortada al 5%	.	.
		Mediana	48,6100	.
		Varianza	155,510	.
		Desv. típ.	12,47035	.
		Mínimo	35,27	.
		Máximo	60,19	.
		Rango	24,92	.
		Amplitud intercuartil	.	.
		Asimetría	-,211	1,225
		Curtosis	.	.
	Positiva	Media	55,1220	3,35111
		Intervalo de confianza para la media al 95%	Límite inferior	48,1080
			Límite superior	62,1360
		Media recortada al 5%	55,2283	.
		Mediana	54,6850	.
		Varianza	224,599	.
		Desv. típ.	14,98662	.
		Mínimo	23,33	.
		Máximo	85,00	.
		Rango	61,67	.
		Amplitud intercuartil	19,66	.
		Asimetría	-,115	,512
		Curtosis	-,004	,992

3.9.3 DIAGRAMA DE CAJAS: TUS



3.9.4 RESULTADOS

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
Quality	Se han asumido varianzas iguales	,496	,489	-,776	21	,446	-7,09867	9,14188	-26,11024	11,91291
	No se han asumido varianzas iguales			-,894	2,946	,438	-7,09867	7,94144	-32,63660	18,43927

Dada la diferencia de medias a variable experiencia académica influye sobre la variable respuesta: Calidad.

3.10 CALIDAD – EXPERIENCIA JAVA

3.10.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Progr_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,323	3	.	,879	3	,322
	Positiva	,129	20	,200*	,979	20	,927

*. Este es un límite inferior de la significación verdadera.

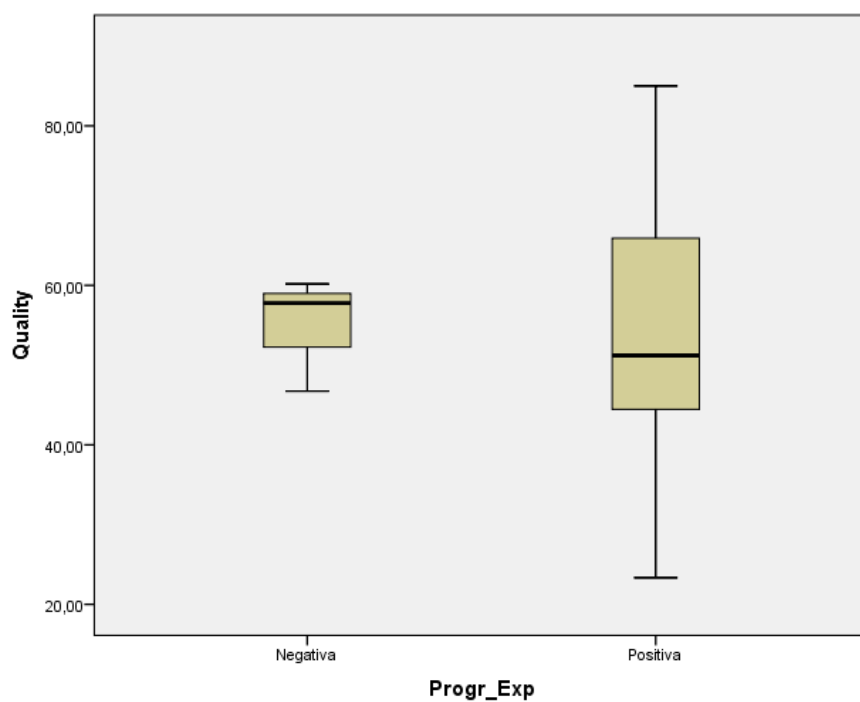
a. Corrección de la significación de Lilliefors

La variable Experiencia de programación supera el test de normalidad.

3.10.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Progr. Exp			Estadístico	Error típ.	
Quality	Negativa	Media	54,8967	4,14710	
		Intervalo de confianza para la media al 95%	Límite inferior	37,0531	
			Límite superior	72,7402	
		Media recortada al 5%	.		
		Mediana	57,7800		
		Varianza	51,595		
		Desv. típ.	7,18300		
		Mínimo	46,72		
		Máximo	60,19		
		Rango	13,47		
		Amplitud intercuartil	.		
		Asimetría	-1,515	1,225	
		Curtosis	.		
	Positiva	Media	54,0910	3,48114	
		Intervalo de confianza para la media al 95%	Límite inferior	46,8049	
			Límite superior	61,3771	
		Media recortada al 5%	54,0828		
		Mediana	51,2050		
		Varianza	242,367		
		Desv. típ.	15,56813		
		Mínimo	23,33		
		Máximo	85,00		
		Rango	61,67		
		Amplitud intercuartil	23,05		
		Asimetría	,002	,512	
		Curtosis	-,352	,992	

3.10.3 DIAGRAMA DE CAJAS: TUS



3.10.4 RESULTADOS

Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia
Quality	Se han asumido varianzas iguales	2,155	,157	,087	21	,932	,80567	9,27052	-18,47344 20,08478
	No se han asumido varianzas iguales			,149	5,523	,887	,80567	5,41450	-12,72564 14,33697

Dado que la diferencia de medias no es significativa no se puede afirmar que la variable Experiencia de programación influye sobre la variable respuesta: Calidad.

3.11 CALIDAD – EXPERIENCIA EN TDD

3.11.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	TDD_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,126	11	,200 [*]	,960	11	,773
	Positiva	,138	12	,200 [*]	,979	12	,979

*. Este es un límite inferior de la significación verdadera.

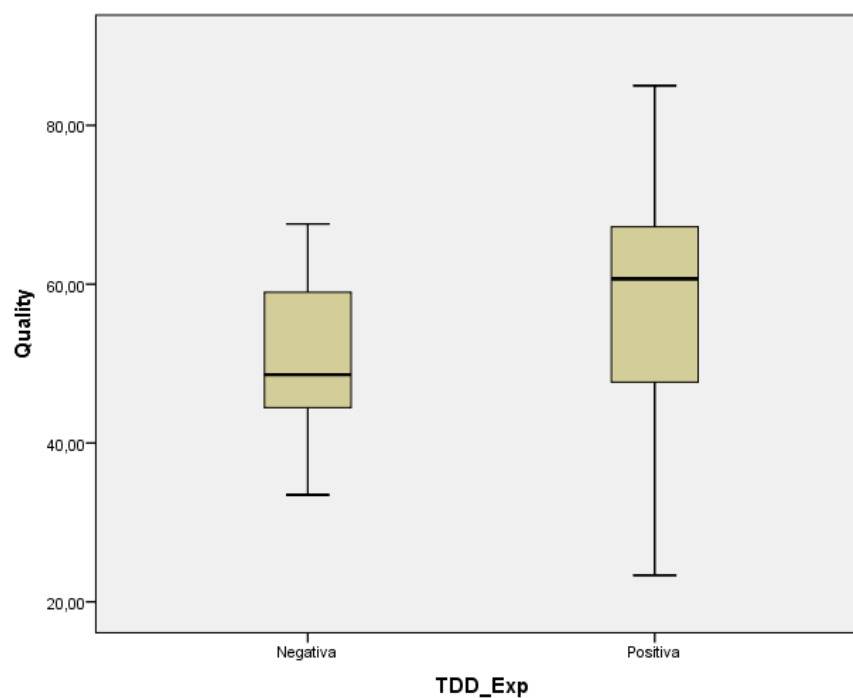
a. Corrección de la significación de Lilliefors

La variable experiencia en TDD supera el test de normalidad.

3.11.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
TDD_Exp			Estadístico	Error típ.
Quality	Negativa	Media	50,5164	3,36358
		Intervalo de confianza para la media al 95%	Límite inferior	43,0218
			Límite superior	58,0109
		Media recortada al 5%	50,5143	
		Mediana	48,6100	
		Varianza	124,451	
		Desv. típ.	11,15574	
		Mínimo	33,48	
		Máximo	67,59	
		Rango	34,11	
		Amplitud intercuartil	17,99	
		Asimetría	-,016	,661
		Curtosis	-,887	1,279
	Positiva	Media	57,5692	4,90815
		Intervalo de confianza para la media al 95%	Límite inferior	46,7664
			Límite superior	68,3719
		Media recortada al 5%	57,9474	
		Mediana	60,6850	
		Varianza	289,079	
		Desv. típ.	17,00233	
		Mínimo	23,33	
		Máximo	85,00	
		Rango	61,67	
		Amplitud intercuartil	20,21	
		Asimetría	-,437	,637
		Curtosis	,229	1,232

3.11.3 DIAGRAMA DE CAJAS: TUS



Resultados

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% intervalo de confianza para la diferencia	
Quality	Se han asumido varianzas iguales	1,644	,214	-1,164	21	,257	-7,05280	6,05889	-19,65296	5,54736
	No se han asumido varianzas iguales			-1,185	19,120	,250	-7,05280	5,95009	-19,50123	5,39562

La variable respuesta Experiencia en TDD influye sobre la variable respuesta: Calidad.

3.12 CALIDAD – EXPERIENCIA EN JUNIT

3.12.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	JUnit_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,213	8	,200*	,868	8	,143
	Positiva	,124	15	,200*	,983	15	,985

*. Este es un límite inferior de la significación verdadera.

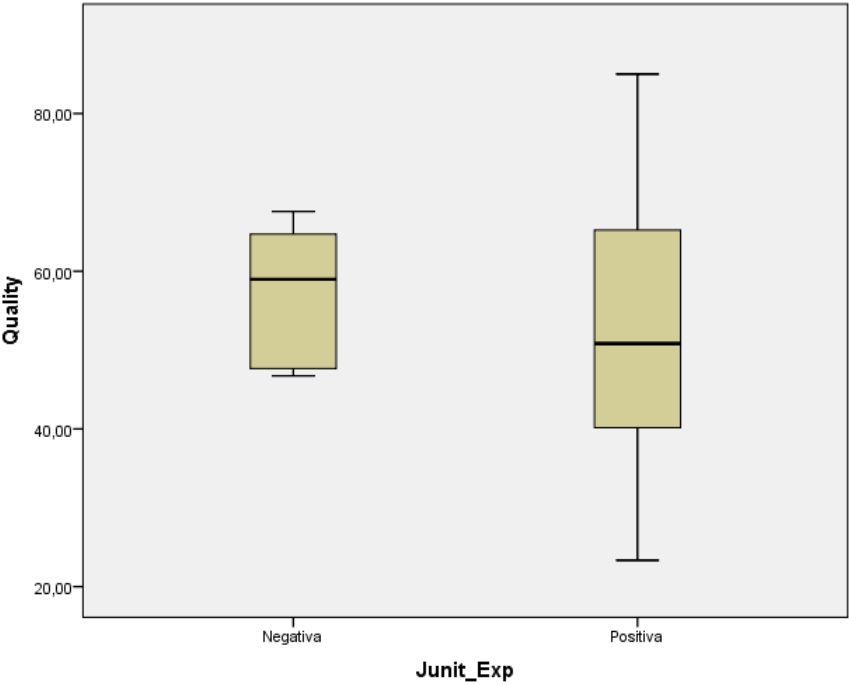
a. Corrección de la significación de Lilliefors

La variable Experiencia en JUnit supera el test de normalidad.

3.12.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Junit_Exp			Estadístico	Error tip.
Quality	Negativa	Media	57,1275	3,05737
		Intervalo de confianza para la media al 95%	Límite inferior	49,8980
			Límite superior	64,3570
		Media recortada al 5%	57,1244	
		Mediana	58,9850	
		Varianza	74,780	
		Desv. tip.	8,64754	
		Mínimo	46,72	
		Máximo	67,59	
		Rango	20,87	
		Amplitud intercuartil	17,66	
		Asimetría	-,243	,752
		Curtosis	-1,991	1,481
	Positiva	Media	52,6327	4,40832
		Intervalo de confianza para la media al 95%	Límite inferior	43,1778
			Límite superior	62,0876
		Media recortada al 5%	52,4624	
		Mediana	50,8200	
		Varianza	291,499	
		Desv. tip.	17,07334	
		Mínimo	23,33	
		Máximo	85,00	
		Rango	61,67	
		Amplitud intercuartil	28,73	
		Asimetría	,227	,580
		Curtosis	-,463	1,121

3.12.3 DIAGRAMA DE CAJAS: TUS



3.12.4 RESULTADOS

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
Quality	Se han asumido varianzas iguales	2,994	,098	,693	21	,496	4,49483	6,48264	-8,98656	17,97623
	No se han asumido varianzas iguales			,838	20,993	,412	4,49483	5,36477	-6,66204	15,65171

La variable Experiencia en JUnit influye en la variable respuesta Calidad.

3.13 CALIDAD – EXPERIENCIA IDE

3.13.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	IDE_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,159	4	.	,993	4	,971
	Positiva	,135	19	,200 [*]	,971	19	,797

*. Este es un límite inferior de la significación verdadera.

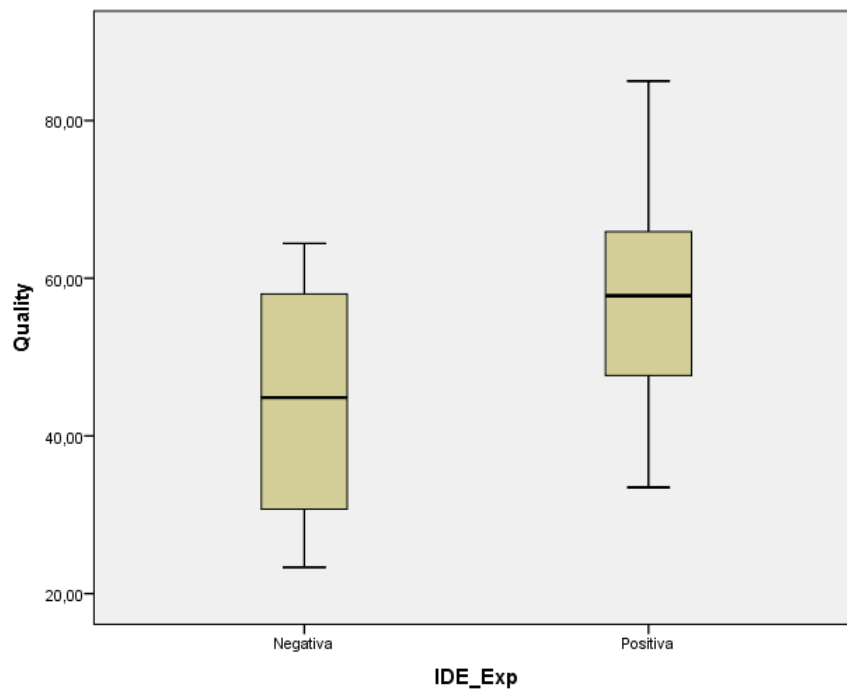
a. Corrección de la significación de Lilliefors

La variable Experiencia en IDEs supera el test de normalidad.

3.13.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
IDE Exp			Estadístico	Error típ.	
Quality	Negativa	Media	44,3600	8,82987	
		Intervalo de confianza para la media al 95%	Límite inferior	16,2594	
			Límite superior	72,4606	
		Media recortada al 5%	44,4144		
		Mediana	44,8500		
		Varianza	311,866		
		Desv. típ.	17,65974		
		Mínimo	23,33		
		Máximo	64,41		
		Rango	41,08		
		Amplitud intercuartil	34,18		
		Asimetría	-,134	1,014	
		Curtosis	-1,112	2,619	
	Positiva	Media	56,2668	3,10903	
		Intervalo de confianza para la media al 95%	Límite inferior	49,7350	
			Límite superior	62,7987	
		Media recortada al 5%	55,9365		
		Mediana	57,7800		
		Varianza	183,655		
		Desv. típ.	13,55193		
		Mínimo	33,48		
		Máximo	85,00		
		Rango	51,52		
		Amplitud intercuartil	20,12		
		Asimetría	,257	,524	
		Curtosis	-,260	1,014	

3.13.3 DIAGRAMA DE CAJAS: TUS



3.13.4 RESULTADOS

Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia
Quality	Se han asumido varianzas iguales	,389	,540	-1,523	21	,143	-11,90684	7,81810	-28,16548 4,35179
	No se han asumido varianzas iguales			-1,272	3,780	,276	-11,90684	9,36123	-38,50451 14,69083

Dada la diferencia de medias, la variable Experiencia en el uso de IDEs influye en la variable respuesta: Calidad.

3.14 CALIDAD – EXPERIENCIA PRUEBAS DE UNIDAD

3.14.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Unit_Test_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,159	7	,200 [*]	,948	7	,710
	Positiva	,123	16	,200 [*]	,972	16	,867

*. Este es un límite inferior de la significación verdadera.

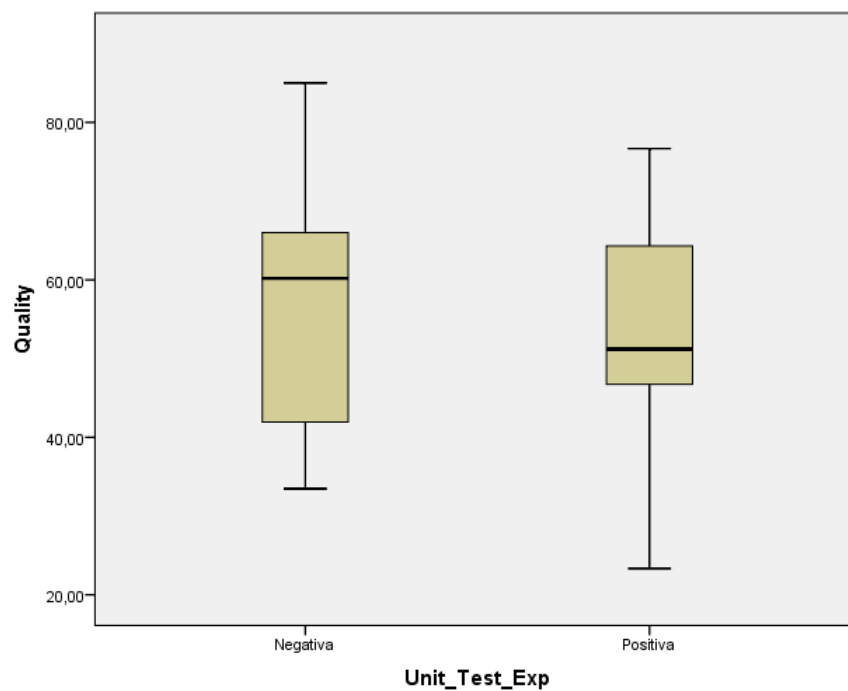
a. Corrección de la significación de Lilliefors

La variable Experiencia en pruebas de unidad supera el test de normalidad.

3.14.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Unit Test Exp			Estadístico	Error tip.
Quality	Negativa	Media	56,3643	6,99350
		Intervalo de confianza para la media al 95%	Límite inferior	39,2518
			Límite superior	73,4768
		Media recortada al 5%	56,0448	
		Mediana	60,1900	
		Varianza	342,363	
		Desv. tip.	18,50306	
		Mínimo	33,48	
		Máximo	85,00	
		Rango	51,52	
		Amplitud intercuartil	32,32	
		Asimetría	,134	,794
		Curtosis	-,748	1,587
	Positiva	Media	53,2475	3,29670
		Intervalo de confianza para la media al 95%	Límite inferior	46,2208
			Límite superior	60,2742
		Media recortada al 5%	53,6083	
		Mediana	51,2050	
		Varianza	173,891	
		Desv. tip.	13,18679	
		Mínimo	23,33	
		Máximo	76,67	
		Rango	53,34	
		Amplitud intercuartil	17,93	
		Asimetría	-,370	,564
		Curtosis	,515	1,091

3.14.3 DIAGRAMA DE CAJAS:



3.14.4 RESULTADOS

Prueba de muestras independientes										
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
Quality	Se han asumido varianzas iguales	1,425	,246	,462	21	,649	3,11679	6,75238	-10,92556	17,15913
	No se han asumido varianzas iguales			,403	8,789	,696	3,11679	7,73157	-14,43739	20,67097

La variable Experiencia en pruebas de unidad influye en la variable respuesta: Calidad.

3.15 CALIDAD – EXPERIENCIA DE PROGRAMACIÓN

3.15.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad ^a							
	Programacion_Exp	Kolmogorov-Smirnov ^b			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,197	4	.	,974	4	,867
	Positiva	,102	19	,200*	,987	19	,994

*. Este es un límite inferior de la significación verdadera.

a. No hay ningún caso válido para Quality cuando Programacion_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

b. Corrección de la significación de Lilliefors

La variable Experiencia de programación supera el test de normalidad.

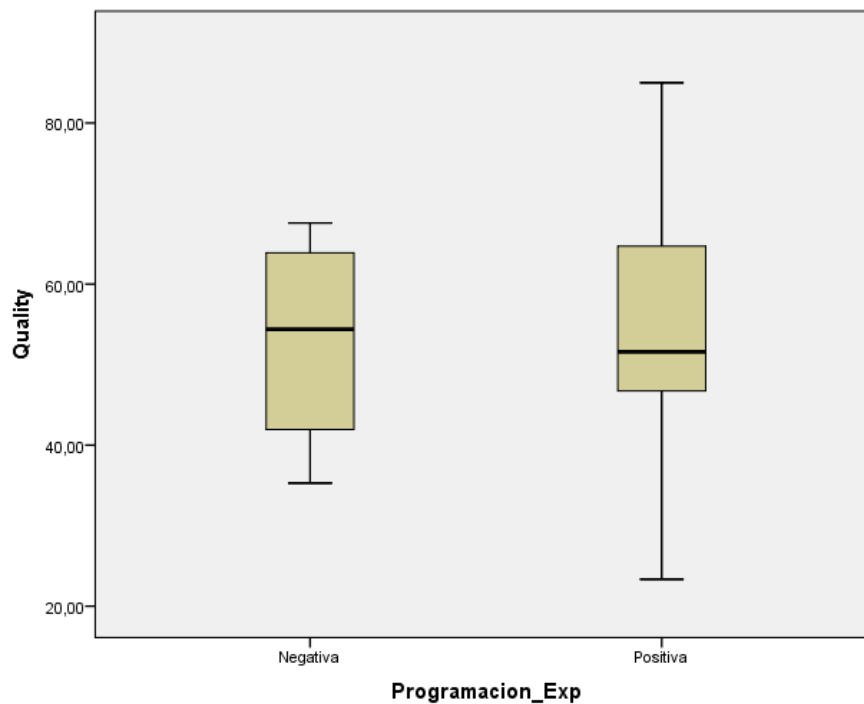
3.15.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos^a

Programacion_Exp				Estadístico	Error típ.
Quality	Negativa	Media		52,9150	7,06022
		Intervalo de confianza para la media al 95%	Límite inferior	30,4462	
			Límite superior	75,3838	
		Media recortada al 5%		53,0800	
		Mediana		54,4000	
		Varianza		199,387	
		Desv. típ.		14,12044	
		Mínimo		35,27	
		Máximo		67,59	
		Rango		32,32	
		Amplitud intercuartil		27,14	
		Asimetría		-,480	1,014
		Curtosis		-1,220	2,619
	Positiva	Media		54,4658	3,46399
		Intervalo de confianza para la media al 95%	Límite inferior	47,1882	
			Límite superior	61,7434	
		Media recortada al 5%		54,4992	
		Mediana		51,5900	
		Varianza		227,986	
		Desv. típ.		15,09920	
		Mínimo		23,33	
		Máximo		85,00	
		Rango		61,67	
		Amplitud intercuartil		18,28	
		Asimetría		-,009	,524
		Curtosis		,059	1,014

a. No hay ningún caso válido para Quality cuando Programacion_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

3.15.3 DIAGRAMA DE CAJAS



3.15.4 RESULTADOS

Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia
Quality	Se han asumido varianzas iguales	,055	,817	-,188	21	,852	-1,55079	8,23160	-18,66935 15,56777
	No se han asumido varianzas iguales			-,197	4,574	,852	-1,55079	7,86422	-22,34604 19,24446

3.16 CALIDAD – METODOLOGÍAS DE DESARROLLO

3.16.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad ^a							
	MetDes_Exp	Kolmogorov-Smirnov ^b			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Quality	Negativa	,202	4	.	,981	4	,909
	Positiva	,100	19	,200*	,987	19	,993

*. Este es un límite inferior de la significación verdadera.

a. No hay ningún caso válido para Quality cuando MetDes_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

b. Corrección de la significación de Lilliefors

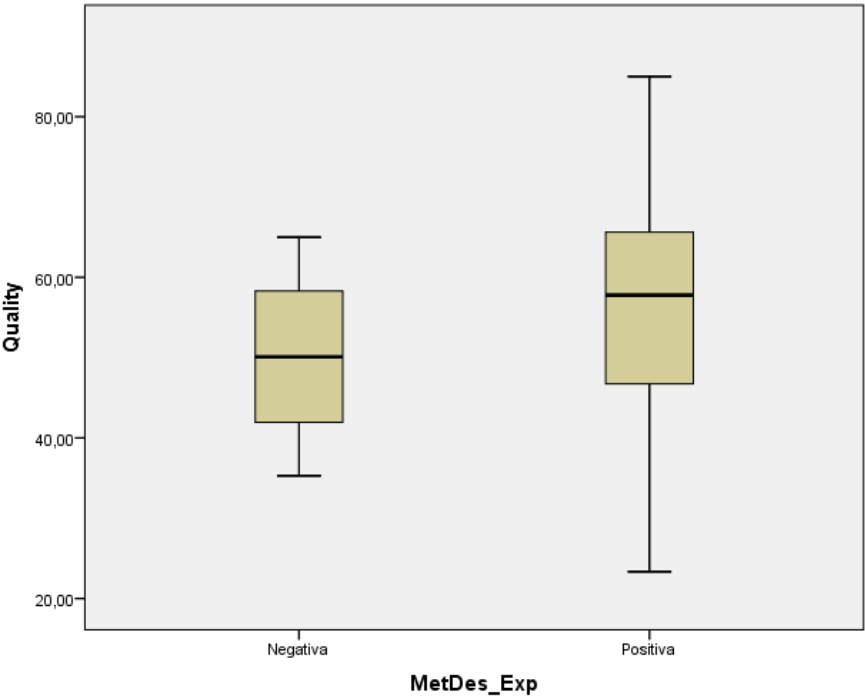
La variable metodologías de desarrollo supera el test de normalidad.

3.16.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos ^a				
MetDes_Exp			Estadístico	Error típ.
Quality	Negativa	Media	50,1175	6,09903
		Intervalo de confianza para la media al 95%	Límite inferior	30,7077
			Límite superior	69,5273
		Media recortada al 5%	50,1156	
		Mediana	50,1000	
		Varianza	148,793	
		Desv. típ.	12,19806	
		Mínimo	35,27	
		Máximo	65,00	
		Rango	29,73	
		Amplitud intercuartil	23,04	
		Asimetría	,008	1,014
		Curtosis	1,205	2,619
	Positiva	Media	55,0547	3,49725
		Intervalo de confianza para la media al 95%	Límite inferior	47,7073
			Límite superior	62,4022
		Media recortada al 5%	55,1536	
		Mediana	57,7800	
		Varianza	232,385	
		Desv. típ.	15,24417	
		Mínimo	23,33	
		Máximo	85,00	
		Rango	61,67	
		Amplitud intercuartil	20,12	
		Asimetría	-,115	,524
		Curtosis	-,041	1,014

a. No hay ningún caso válido para Quality cuando MetDes_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

3.16.3 DIAGRAMA DE CAJAS



3.16.4 RESULTADOS

Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	95% Intervalo de confianza para la diferencia
Quality	Se han asumido varianzas iguales	,796	,382	-,604	21	,552	-4,93724	8,16781	-21,92312 12,04864
	No se han asumido varianzas iguales			-,702	5,203	,513	-4,93724	7,03057	-22,79948 12,92501

3.17 TUS – EXPERIENCIA ACADÉMICA

3.17.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Academic_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,385	3	.	,750	3	,000
	Positiva	,168	20	,141	,866	20	,010

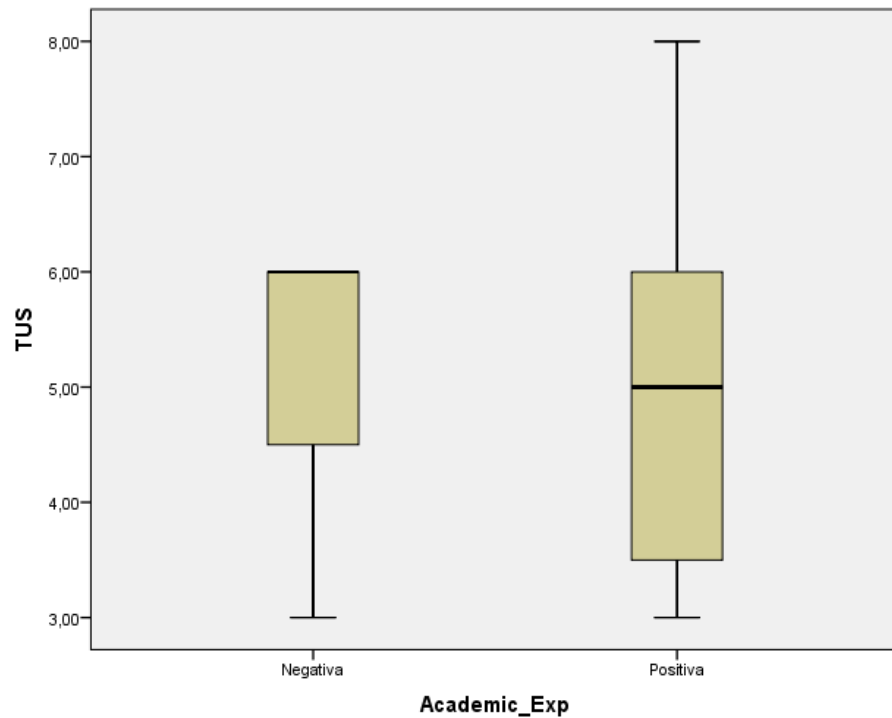
a. Corrección de la significación de Lilliefors

La variable Experiencia académica no supera el test de normalidad.

3.17.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Academic Exp			Estadístico	Error típ.	
TUS	Negativa	Media		5,0000	1,00000
		Intervalo de confianza para la media al 95%	Límite inferior	,6973	
			Límite superior	9,3027	
		Media recortada al 5%		.	
		Mediana		6,0000	
		Varianza		3,000	
		Desv. típ.		1,73205	
		Mínimo		3,00	
		Máximo		6,00	
		Rango		3,00	
		Amplitud intercuartil		.	
		Asimetría		-1,732	1,225
		Curtosis		.	.
	Positiva	Media		5,0500	,40701
		Intervalo de confianza para la media al 95%	Límite inferior	4,1981	
			Límite superior	5,9019	
		Media recortada al 5%		5,0000	
		Mediana		5,0000	
		Varianza		3,313	
		Desv. típ.		1,82021	
		Mínimo		3,00	
		Máximo		8,00	
		Rango		5,00	
		Amplitud intercuartil		2,75	
		Asimetría		,558	,512
		Curtosis		-,916	,992

3.17.3 DIAGRAMA DE CAJAS: TUS



3.17.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Academic_Exp.	Prueba U de Mann-Whitney de muestras independientes	,898 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La diferencia de medias no es significativa como para afirmar que la variable Experiencia Académica influye sobre la variable respuesta: TUS.

3.18 TUS – EXPERIENCIA JAVA

3.18.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	Progr_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,385	3	.	,750	3	,000

Positiva	,147	20	,200*	,900	20	,042
----------	------	----	-------	------	----	------

*. Este es un límite inferior de la significación verdadera.

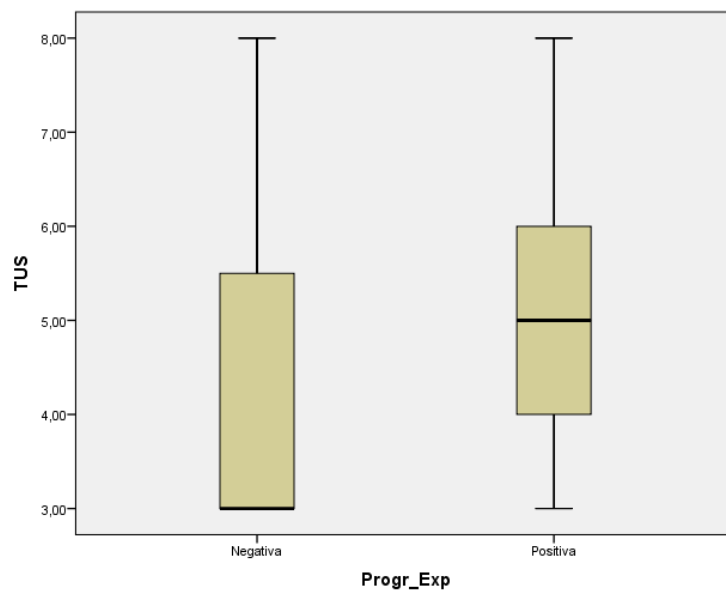
a. Corrección de la significación de Lilliefors

La variable Experiencia de programación no supera el test de normalidad

3.18.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos					
Progr_Exp				Estadístico	Error típ.
TUS	Negativa	Media		4,6667	1,66667
		Intervalo de confianza para la media al 95%	Límite inferior	-2,5044	
			Límite superior	11,8378	
		Media recortada al 5%		.	
		Mediana		3,0000	
		Varianza		8,333	
		Desv. típ.		2,88675	
		Mínimo		3,00	
		Máximo		8,00	
		Rango		5,00	
		Amplitud intercuartil		.	
		Asimetría		1,732	1,225
		Curtosis		.	.
	Positiva	Media		5,1000	,36921
		Intervalo de confianza para la media al 95%	Límite inferior	4,3272	
			Límite superior	5,8728	
		Media recortada al 5%		5,0556	
		Mediana		5,0000	
		Varianza		2,726	
		Desv. típ.		1,65116	
		Mínimo		3,00	
		Máximo		8,00	
		Rango		5,00	
		Amplitud intercuartil		2,00	
		Asimetría		,447	,512
		Curtosis		-,632	,992

3.18.3 DIAGRAMA DE CAJAS: TUS



3.18.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Progr_Exp.	Prueba U de Mann-Whitney de muestras independientes	,514 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Dado que la diferencia de medias no es significativa, no se pudo asegurar que la variable Experiencia de programación influye en la variable respuesta: TUS.

3.19 TUS – EXPERIENCIA EN TDD

3.19.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	TDD_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,170	11	,200*	,889	11	,135
	Positiva	,262	12	,023	,852	12	,039

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

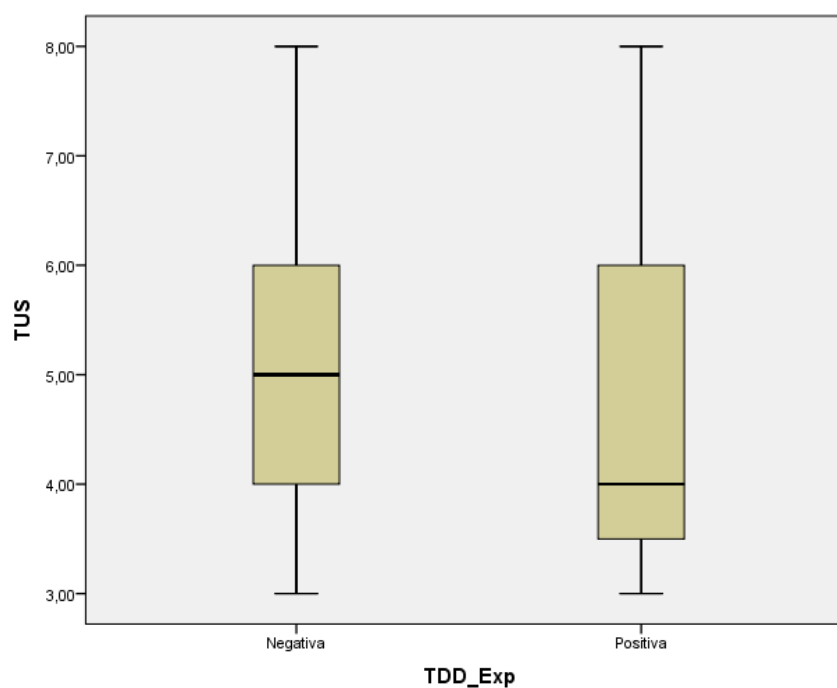
La variable Experiencia en TDD no supera el test de normalidad.

3.19.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

TDD_Exp				Estadístico	Error típ.
TUS	Negativa	Media		5,2727	,54089
		Intervalo de confianza para la media al 95%	Límite inferior	4,0675	
			Límite superior	6,4779	
		Media recortada al 5%		5,2475	
		Mediana		5,0000	
		Varianza		3,218	
		Desv. típ.		1,79393	
		Mínimo		3,00	
		Máximo		8,00	
		Rango		5,00	
		Amplitud intercuartil		3,00	
		Asimetría		,136	,661
		Curtosis		-,776	1,279
	Positiva	Media		4,8333	,51981
		Intervalo de confianza para la media al 95%	Límite inferior	3,6892	
			Límite superior	5,9774	
		Media recortada al 5%		4,7593	
		Mediana		4,0000	
		Varianza		3,242	
		Desv. típ.		1,80067	
		Mínimo		3,00	
		Máximo		8,00	
		Rango		5,00	
		Amplitud intercuartil		2,75	
		Asimetría		,857	,637
		Curtosis		-,427	1,232

3.19.3 DIAGRAMA DE CAJAS: TUS



3.19.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de TDD_Exp.	Prueba U de Mann-Whitney de muestras independientes	,566 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Dado que la diferencia de medias no es significativa, no se puede asegurar que la variable Experiencia en TDD influye en la variable respuesta: TUS.

3.20 TUS – EXPERIENCIA EN JUNIT

3.20.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Junit_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,200	8	,200*	,841	8	,078
	Positiva	,149	15	,200*	,911	15	,139

*. Este es un límite inferior de la significación verdadera.

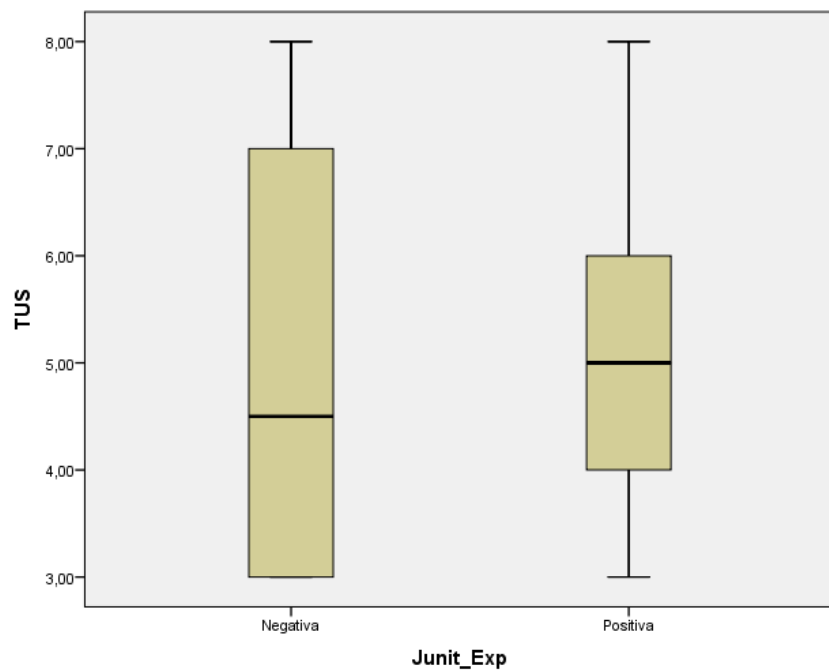
a. Corrección de la significación de Lilliefors

La variable Experiencia en JUnit supera el test de normalidad.

3.20.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
Junit_Exp			Estadístico	Error típ.
TUS	Negativa	Media	5,0000	,75593
		Intervalo de confianza para la media al 95%	Límite inferior 3,2125	
			Límite superior 6,7875	
		Media recortada al 5%	4,9444	
		Mediana	4,5000	
		Varianza	4,571	
		Desv. típ.	2,13809	
		Mínimo	3,00	
		Máximo	8,00	
		Rango	5,00	
		Amplitud intercuartil	4,50	
		Asimetría	,585	,752
		Curtosis	-1,422	1,481
	Positiva	Media	5,0667	,41937
		Intervalo de confianza para la media al 95%	Límite inferior 4,1672	
			Límite superior 5,9661	
		Media recortada al 5%	5,0185	
		Mediana	5,0000	
		Varianza	2,638	
		Desv. típ.	1,62422	
		Mínimo	3,00	
		Máximo	8,00	
		Rango	5,00	
		Amplitud intercuartil	2,00	
		Asimetría	,454	,580
		Curtosis	-,458	1,121

3.20.3 DIAGRAMA DE CAJAS: TUS



3.20.4 RESULTADOS

Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias					
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia
TUS	Se han asumido varianzas iguales	1,219	,282	-,084	21	,934	-,06667	,79319	-1,71620 1,58287
	No se han asumido varianzas iguales			-,077	11,431	,940	-,06667	,86447	-1,96064 1,82730

Dada que la diferencia de medias no es significativa, no se puede asegurar que la variable Experiencia en JUnit influye en la variable respuesta: TUS.

3.21 TUS – EXPERIENCIA IDE

3.21.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad							
	IDE_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,298	4	.	,926	4	,572
	Positiva	,188	19	,075	,876	19	,018

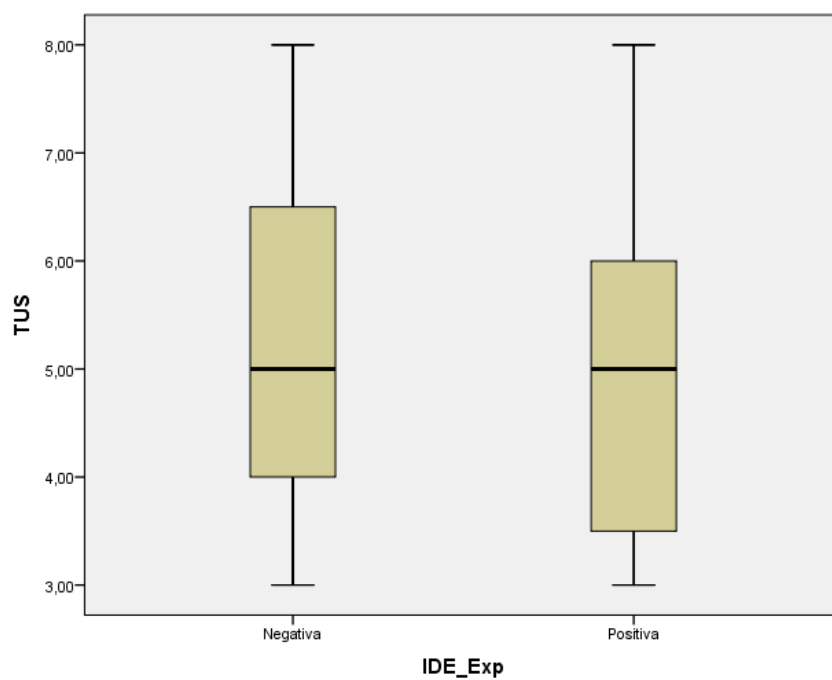
a. Corrección de la significación de Lilliefors

La variable experiencia en el uso de IDEs no supera el test de normalidad.

3.21.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos				
IDE_Exp			Estadístico	Error típ.
TUS	Negativa	Media	5,2500	1,03078
		Intervalo de confianza para la media al 95%	Límite inferior	1,9696
			Límite superior	8,5304
		Media recortada al 5%	5,2222	
		Mediana	5,0000	
		Varianza	4,250	
		Desv. típ.	2,06155	
		Mínimo	3,00	
		Máximo	8,00	
		Rango	5,00	
		Amplitud intercuartil	3,75	
		Asimetría	,713	1,014
		Curtosis	1,785	2,619
	Positiva	Media	5,0000	,40465
		Intervalo de confianza para la media al 95%	Límite inferior	4,1499
			Límite superior	5,8501
		Media recortada al 5%	4,9444	
		Mediana	5,0000	
		Varianza	3,111	
		Desv. típ.	1,76383	
		Mínimo	3,00	
		Máximo	8,00	
		Rango	5,00	
		Amplitud intercuartil	3,00	
		Asimetría	,475	,524
		Curtosis	-,911	1,014

3.21.3 DIAGRAMA DE CAJAS: TUS



3.21.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de IDE_Exp.	Prueba U de Mann-Whitney de muestras independientes	,907 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

Dado que la diferencia de medias no es significativa, no se puede asegurar que la variable Experiencia en el uso de IDEs influye en la variable respuesta: TUS.

3.22 TUS – EXPERIENCIA PRUEBAS DE UNIDAD

3.22.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad

	Unit_Test_Exp	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,203	7	,200 [*]	,877	7	,215
	Positiva	,177	16	,195	,863	16	,021

*. Este es un límite inferior de la significación verdadera.

a. Corrección de la significación de Lilliefors

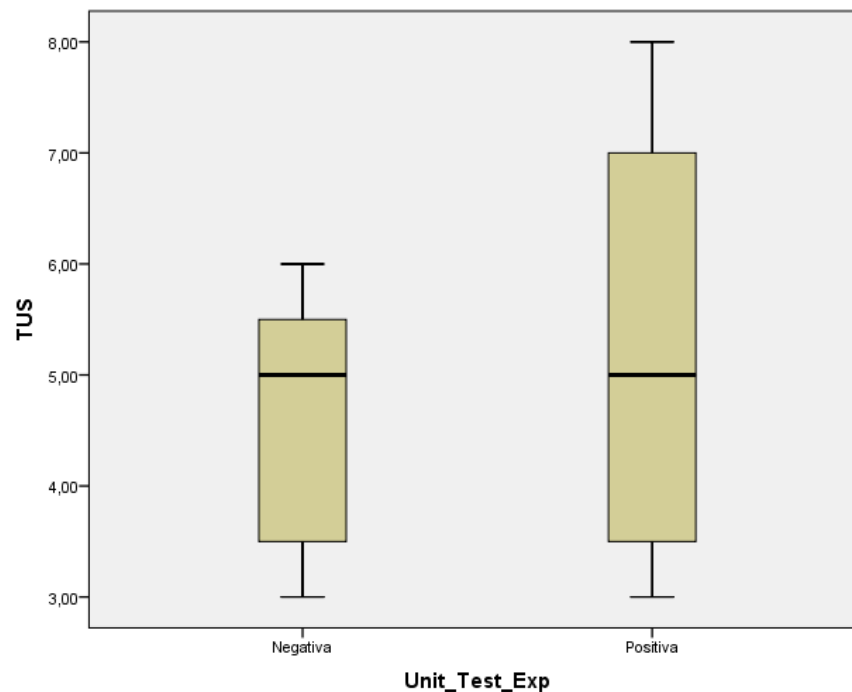
La variable Experiencia en pruebas de unidad no supera el test de normalidad.

3.22.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos

Unit Test Exp				Estadístico	Error típ.
TUS	Negativa	Media		4,5714	,48093
		Intervalo de confianza para la media al 95%	Límite inferior	3,3946	
			Límite superior	5,7482	
		Media recortada al 5%		4,5794	
		Mediana		5,0000	
		Varianza		1,619	
		Desv. típ.		1,27242	
		Mínimo		3,00	
		Máximo		6,00	
		Rango		3,00	
		Amplitud intercuartil		3,00	
		Asimetría		-,222	,794
		Curtosis		-1,715	1,587
	Positiva	Media		5,2500	,48734
		Intervalo de confianza para la media al 95%	Límite inferior	4,2113	
			Límite superior	6,2887	
		Media recortada al 5%		5,2222	
		Mediana		5,0000	
		Varianza		3,800	
		Desv. típ.		1,94936	
		Mínimo		3,00	
		Máximo		8,00	
		Rango		5,00	
		Amplitud intercuartil		4,25	
		Asimetría		,339	,564
		Curtosis		-1,366	1,091

3.22.3 DIAGRAMA DE CAJAS: TUS



3.22.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Unit_Test_Exp.	Prueba U de Mann-Whitney de muestras independientes	,535 ¹	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

¹Se muestra la significancia exacta para esta prueba.

La diferencia de medias no es significativa, no se puede asegurar que la variable Experiencia en pruebas de unidad influye en la variable respuesta: TUS.

3.23 TUS – EXPERIENCIA DE PROGRAMACIÓN

3.23.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad^a

	Programacion_Exp	Kolmogorov-Smirnov ^b			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,307	4	.	,729	4	,024
	Positiva	,166	19	,175	,876	19	,018

a. No hay ningún caso válido para TUS cuando Programacion_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

b. Corrección de la significación de Lilliefors

La variable experiencia de programación no supera el test de normalidad.

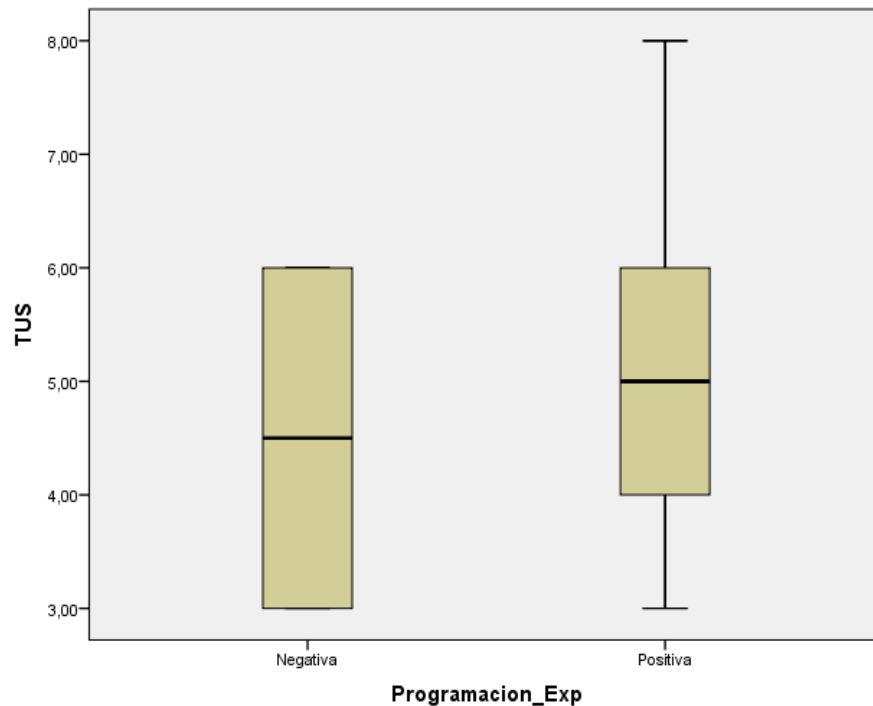
3.23.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos^a

Programacion_Exp			Estadístico	Error típ.
TUS	Negativa	Media	4,5000	,86603
		Intervalo de confianza para la media al 95%	Límite inferior	1,7439
			Límite superior	7,2561
		Media recortada al 5%	4,5000	
		Mediana	4,5000	
		Varianza	3,000	
		Desv. típ.	1,73205	
		Mínimo	3,00	
		Máximo	6,00	
		Rango	3,00	
		Amplitud intercuartil	3,00	
		Asimetría	,000	1,014
		Curtosis	-6,000	2,619
	Positiva	Media	5,1579	,41368
		Intervalo de confianza para la media al 95%	Límite inferior	4,2888
			Límite superior	6,0270
		Media recortada al 5%	5,1199	
		Mediana	5,0000	
		Varianza	3,251	
		Desv. típ.	1,80318	
		Mínimo	3,00	
		Máximo	8,00	
		Rango	5,00	
		Amplitud intercuartil	2,00	
		Asimetría	,500	,524
		Curtosis	-,957	1,014

a. No hay ningún caso válido para TUS cuando Programacion_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

3.23.3 DIAGRAMA DE CAJAS



3.23.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de Programacion_Exp.	Prueba Kruskal-Wallis de muestras independientes	,562	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

Ha sido necesario aplicar la Pruebas de Kruskal-Wallis, puesto que Man Withney no puede calcular....

3.24 TUS – METODOLOGÍAS DE DESARROLLO

3.24.1 PRUEBA DE NORMALIDAD

Pruebas de normalidad^a

	MetDes_Exp	Kolmogorov-Smirnov ^b			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
TUS	Negativa	,250	4	.	,945	4	,683
	Positiva	,167	19	,175	,861	19	,010

a. No hay ningún caso válido para TUS cuando MetDes_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

b. Corrección de la significación de Lilliefors

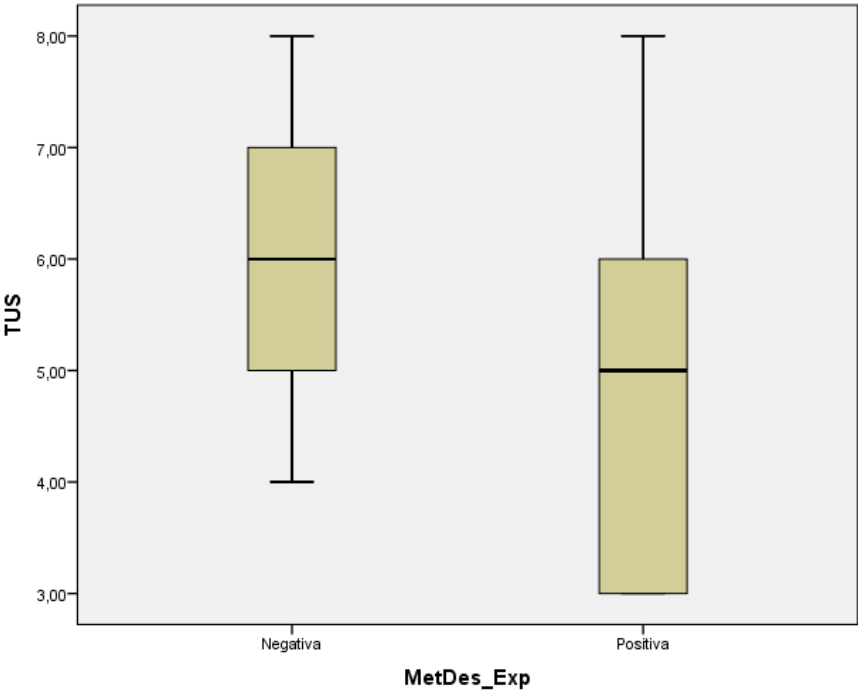
La variable Experiencia en metodologías de desarrollo no supera el test de normalidad.

3.24.2 ESTADÍSTICOS DESCRIPTIVOS

Descriptivos ^a					
MetDes_Exp				Estadístico	Error típ.
TUS	Negativa	Media		6,0000	,81650
		Intervalo de confianza para la media al 95%	Límite inferior	3,4015	
			Límite superior	8,5985	
		Media recortada al 5%		6,0000	
		Mediana		6,0000	
		Varianza		2,667	
		Desv. típ.		1,63299	
		Mínimo		4,00	
		Máximo		8,00	
		Rango		4,00	
		Amplitud intercuartil		3,00	
		Asimetría		,000	1,014
		Curtosis		1,500	2,619
	Positiva	Media		4,8421	,40655
		Intervalo de confianza para la media al 95%	Límite inferior	3,9880	
			Límite superior	5,6962	
		Media recortada al 5%		4,7690	
		Mediana		5,0000	
		Varianza		3,140	
		Desv. típ.		1,77210	
		Mínimo		3,00	
		Máximo		8,00	
		Rango		5,00	
		Amplitud intercuartil		3,00	
		Asimetría		,668	,524
		Curtosis		-,627	1,014

a. No hay ningún caso válido para TUS cuando MetDes_Exp = ,000. No se pueden calcular los estadísticos para este nivel.

3.24.3 DIAGRAMA DE CAJAS



3.24.4 RESULTADOS

Resumen de prueba de hipótesis

	Hipótesis nula	Test	Sig.	Decisión
1	La distribución de TUS es la misma entre las categorías de MetDes_Exp.	Prueba Kruskal-Wallis de muestras independientes	,185	Retener la hipótesis nula.

Se muestran las significancias asintóticas. El nivel de significancia es ,05.

4 RESUMEN DE RESULTADOS DICOTOMIZACIÓN

	ROVER			BOWLING SCOREKEEPER			MUSICPHONE		
	Productividad	Calidad	TUS	Productividad	Calidad	TUS	Productividad	Calidad	TUS
E. Académica	0,135	0,183	0,097	0,002	0,023	0,003	0,451	0,446	0,898
E. Programación	0,241	0,347	0,241	0,005	0,179	0,003	0,744	0,852	0,562
E. Java	0,241	0,273	0,183	0,241	0,737	0,183	0,388	0,932	0,514
E. Testing	0,009	0,052	0,007	0,038	0,045	0,61	0,908	0,649	0,535
E. JUnit	0,599	0,815	0,379	0,318	0,726	0,29	0,5	0,496	0,934
E. IDE	1	0,627	0,911	0,068	0,241	0,135	0,863	0,143	0,907
E. MetDes	0,581	0,945	0,679	0,103	0,16	0,12	0,051	0,552	0,185
E. TDD	0,89	0,178	0,16	0,41	0,045	0,63	0,614	0,257	0,566

Tabla. Resumen de resultados

La tabla tiene el resumen del análisis del apartado anterior, este análisis consiste en comprobar que la variable presenta normalidad, si esto es así se realiza una comparación de medias mediante el análisis de la T Student. Cuando no hay normalidad los resultados que se muestran son los del test de Mann-Withney. Para MusicPhone donde los datos presentan normalidad en su mayoría no hay ninguna media que presente una diferencia significativa. La única variable que resulta influyente según el análisis que se está llevando a cabo es la Experiencia en Testing.

En la ejecución del experimento llevada a cabo en Bowling Scorekeeper la mayoría de variables resultan tener influencia en alguna de las variables respuesta un hecho que resulta curioso y que intentaremos explicar aplicando otros tipos de análisis. Por su parte Mars Rover tampoco presenta un gran grupo de variables influyentes, la única que se repite para las tres variables respuesta en la Experiencia en Testing que ya se mencionó anteriormente.

Es importante recalcar que los resultados obtenidos en la dicotomización, no son muy precisos, esto se debe a que la dicotomización se lleva a cabo según un criterio de división en grupos que tiene un umbral que viene determinado por criterio de quien realiza el análisis. Otro motivo que hace que el análisis dicotómico no sea robusto es el tamaño de la muestra

ANEXO C.

Análisis correlación de Pearson

1 ANÁLISIS CORRELACIÓN MARS ROVERS

Se han cuantificado las variables que en los análisis anteriores las teníamos como dicotómicas para poder hacer un análisis más fiel a los datos obtenidos. A continuación se presenta la correlación de Pearson para cada par de variables.

1.1 PRODUCTIVIDAD - EXPERIENCIA ACADÉMICA

Correlaciones		Productivity	Academica
Productivity	Correlación de Pearson	1	-,175
	Sig. (bilateral)		,413
	N	24	24
Academica	Correlación de Pearson	-,175	1
	Sig. (bilateral)	,413	
	N	24	24

1.2 PRODUCTIVIDAD - EXPERIENCIA DE PROGRAMACIÓN

Correlaciones		Productivity	Programacion
Productivity	Correlación de Pearson	1	-,058
	Sig. (bilateral)		,789
	N	24	24
Programacion	Correlación de Pearson	-,058	1
	Sig. (bilateral)	,789	
	N	24	24

1.3 PRODUCTIVIDAD - EXPERIENCIA EN JAVA

Correlaciones		Productivity	Java
Productivity	Correlación de Pearson	1	,057
	Sig. (bilateral)		,793
	N	24	24
Java	Correlación de Pearson	,057	1
	Sig. (bilateral)	,793	
	N	24	24

1.4 PRODUCTIVIDAD - EXPERIENCIA EN TESTING

Correlaciones		Productivity	Testing
Productivity	Correlación de Pearson	1	,126
	Sig. (bilateral)		,557
	N	24	24
Testing	Correlación de Pearson	,126	1
	Sig. (bilateral)	,557	
	N	24	24

1.5 PRODUCTIVIDAD - EXPERIENCIA EN JUNIT

Correlaciones

		Productivity	JUnit
Productivity	Correlación de Pearson	1	,306
	Sig. (bilateral)		,146
	N	24	24
JUnit	Correlación de Pearson	,306	1
	Sig. (bilateral)	,146	
	N	24	24

1.6 PRODUCTIVIDAD - EXPERIENCIA IDE

Correlaciones

		Productivity	IDE
Productivity	Correlación de Pearson	1	-,061
	Sig. (bilateral)		,777
	N	24	24
IDE	Correlación de Pearson	-,061	1
	Sig. (bilateral)	,777	
	N	24	24

1.7 PRODUCTIVIDAD - EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO

Correlaciones

		Productivity	MetDes
Productivity	Correlación de Pearson	1	-,384
	Sig. (bilateral)		,064
	N	24	24
MetDes	Correlación de Pearson	-,384	1
	Sig. (bilateral)	,064	
	N	24	24

1.8 PRODUCTIVIDAD - EXPERIENCIA EN TDD

Correlaciones

		Productivity	TDD
Productivity	Correlación de Pearson	1	,301
	Sig. (bilateral)		,153
	N	24	24
TDD	Correlación de Pearson	,301	1
	Sig. (bilateral)	,153	
	N	24	24

2 RESUMEN RESULTADOS CORRELACIÓN DE PEARSON

2.1 MARS ROVER

MARS ROVER									
Correlaciones									
		Academica	Programacion	Java	Testing	JUnit	IDE	MetDes	TDD
Productivity	Correlación de Pearson	-,175	-,058	,057	,126	,306	-,061	-,384	,301
	Sig. (bilateral)	,413	,789	,793	,557	,146	,777	,064	,153
	N	24	24	24	24	24	24	24	24
Quality	Correlación de Pearson	-,086	,098	,166	,112	,379	,008	-,116	,409
	Sig. (bilateral)	,690	,650	,438	,601	,068	,969	,589	,047
	N	24	24	24	24	24	24	24	24
TUS	Correlación de Pearson	-,152	-,041	,082	,106	,340	-,113	-,311	,344
	Sig. (bilateral)	,479	,848	,703	,621	,104	,600	,139	,099
	N	24	24	24	24	24	24	24	24

** . La correlación es significativa al nivel 0,01 (bilateral).

* . La correlación es significativa al nivel 0,05 (bilateral).

En términos generales, a la vista de los resultados, las variables que tienen influencia en las variables respuesta son La experiencia en **JUnit**, en **metodologías de desarrollo** y en **TDD**. Dado que solo la Experiencia en TDD frente a la Calidad tiene un coeficiente de correlación significativo, se ha relajado un poco el nivel de significación hasta el 10%.

2.2 BOWLING SCOREKEEPER

BOWLING SCOREKEEPER

Correlaciones

		Academica	Programacion	Java	Testing	JUnit	IDE	MetDes	TDD
Productivity	Correlación de Pearson	,081	,230	,206	,294	,292	-,344	,218	,218
	Sig. (bilateral)	,706	,280	,333	,163	,167	,100	,306	,306
	N	24	24	24	24	24	24	24	24
Quality	Correlación de Pearson	,253	,311	,017	,110	,108	-,228	,342	,360
	Sig. (bilateral)	,232	,140	,936	,608	,616	,284	,102	,084
	N	24	24	24	24	24	24	24	24
TUS	Correlación de Pearson	,283	,398	,423	,471	,471	-,320	,292	,181
	Sig. (bilateral)	,179	,054	,040	,020	,020	,128	,166	,398
	N	24	24	24	24	24	24	24	24

*. La correlación es significativa al nivel 0,05 (bilateral).

**. La correlación es significativa al nivel 0,01 (bilateral).

En el caso de Scorekeeper las variables que presentan influencia en las variables respuesta son la Experiencia en el uso del **IDE** para el caso de la productividad, La experiencia en **metodologías de desarrollo** y en **TDD** para la Calidad, y para el caso de TUS la experiencia de programación, en Java, en **Testing** y en **JUnit**.

2.3 MUSICPHONE

MUSICPHONE									
Correlaciones									
		Academica	Programacion	Java	Testing	JUnit	IDE	MetDes	TDD
Productivity	Correlación de Pearson	-,297	-,249	-,124	-,414	-,082	,065	-,411	-,021
	Sig. (bilateral)	,168	,251	,573	,050	,710	,769	,051	,924
	N	23	23	23	23	23	23	23	23
Quality	Correlación de Pearson	,069	-,155	,051	-,154	-,064	,315	,131	,323
	Sig. (bilateral)	,753	,480	,817	,483	,771	,143	,552	,133
	N	23	23	23	23	23	23	23	23
TUS	Correlación de Pearson	-,162	-,168	-,145	-,198	-,123	-,055	-,253	-,090
	Sig. (bilateral)	,459	,443	,511	,364	,577	,804	,243	,681
	N	23	23	23	23	23	23	23	23

** . La correlación es significativa al nivel 0,01 (bilateral).

* . La correlación es significativa al nivel 0,05 (bilateral).

En este caso las únicas variables que presentan correlación son la Experiencia en **Testing** y la experiencia en **metodologías de desarrollo** frente a la Productividad.

ANEXO D.

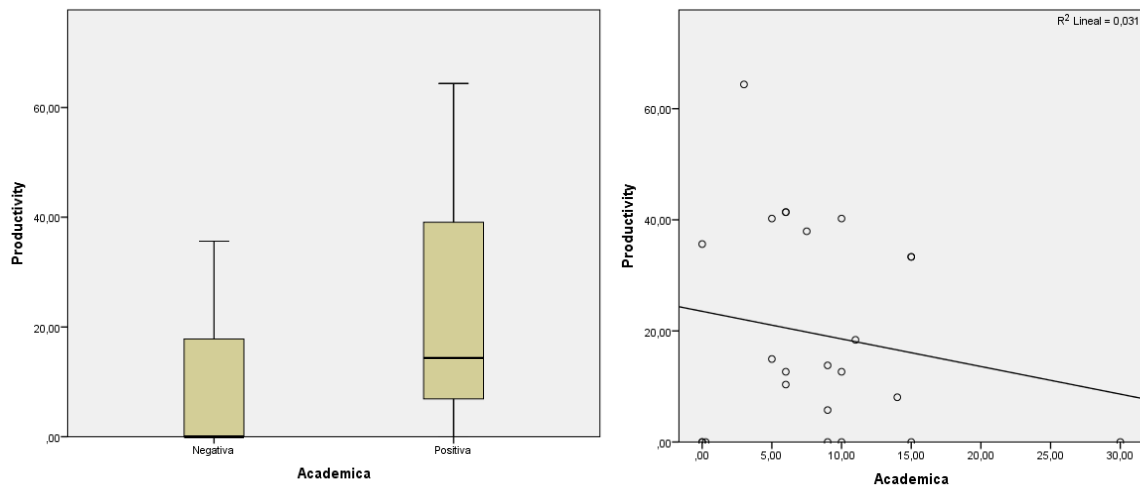
**Comparación de resultados
obtenidos en la Dicotomización
frente a la correlación de Pearson**

1 DICOTOMIZACIÓN Y CORRELACIÓN MARS ROVERS

Este documento contiene un análisis comparativo de los resultados obtenidos en los dos análisis: análisis dicotómico y análisis de correlación de Pearson. En este apartado se presenta el análisis del experimento MarsRover. Igual que en los análisis llevados a cabo anteriormente se compara las variables por pares.

1.1 PRODUCTIVIDAD – EXPERIENCIA ACADÉMICA

A la izquierda de la figura se muestra el diagrama de cajas que se obtiene con los datos del análisis dicotómico, es decir, la clasificación de los sujetos en dos categorías: los que se considera que pueden influir positivamente y los que se considera que pueden influir negativamente. En este caso resulta fácil deducir de la figura que tanto la mediana como la media son superiores para la categoría Positiva. En la derecha de la figura se presenta un gráfico de dispersión con los datos del análisis cuantitativo. Se puede apreciar la información de la experiencia de los sujetos con mucho más detalle y esto permite dilucidar que no hay una relación clara entre la experiencia académica y la productividad. Hay un dato que parece ser un valor atípico de un sujeto con 5 años de experiencia académica que tiene la productividad más alta (64,37) es el causante de que el ajuste lineal resulte ser negativo. Examinando el resto de los datos se concluye que no hay ninguna relación entre la experiencia académica y la productividad.

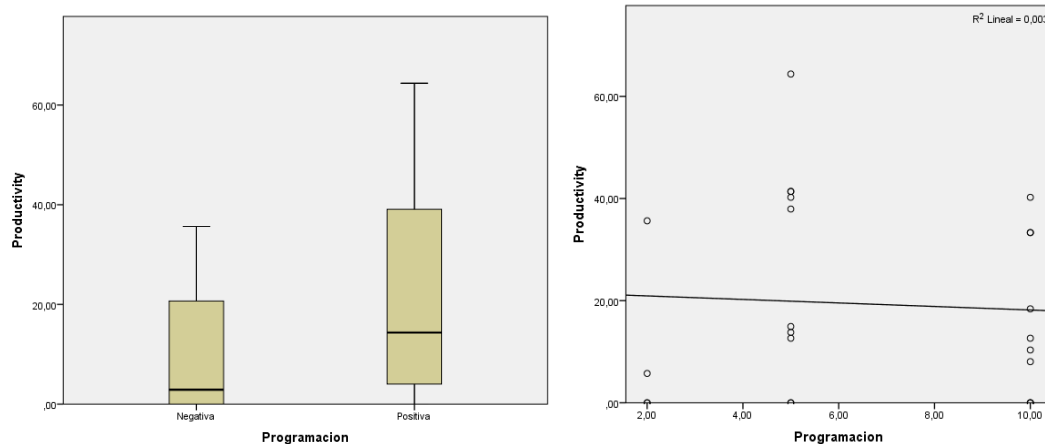


En la parte superior derecha de la figura encontramos la bondad de ajuste de la aproximación lineal que se ve en la figura, es un dato muy bajo (0,031) como para tener en cuenta la pendiente de la recta del ajuste lineal.

El análisis dicotómico muestra una clara influencia de la experiencia académica de los sujetos en la Productividad, sin embargo no se obtiene el mismo resultado del análisis de correlación de Pearson. Con los datos cuantitativos se obtiene una representación más fiel y permite ver con más detalle que es lo que realmente está sucediendo.

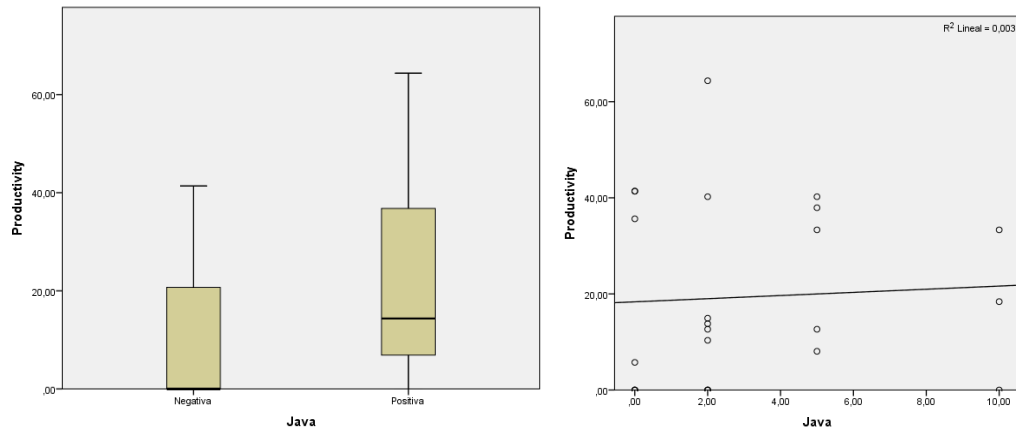
1.2 PRODUCTIVIDAD – EXPERIENCIA DE PROGRAMACIÓN

En la izquierda de la figura se muestra el diagrama de cajas para los datos dicotómicos de la experiencia de programación frente a la Productividad. Del diagrama de cajas se puede deducir poco, solo se puede confirmar que la mediana es superior en la categoría positiva. La grafica de dispersión de la derecha da más luz de lo que está ocurriendo con los datos. La mayoría de los sujetos tienen más de cinco años de experiencia sin embargo del gráfico de dispersión no se puede afirmar por ejemplo que los sujetos con mayor experiencia de programación presenten mayor Productividad puesto que hay muchos datos que contradicen esta afirmación.



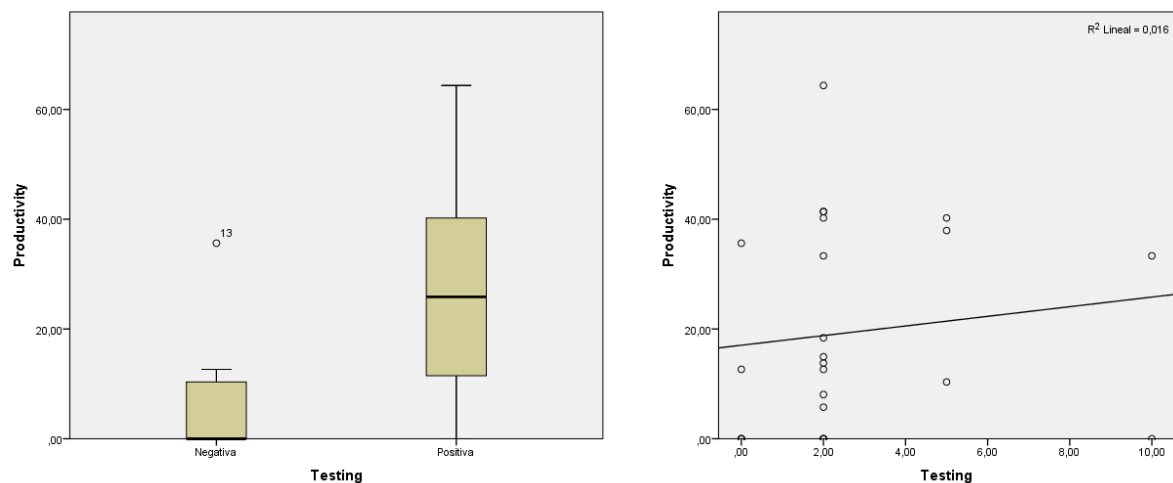
1.3 PRODUCTIVIDAD – EXPERIENCIA EN JAVA

En el diagrama de cajas de la parte izquierda de la figura se aprecia que los sujetos que tienen más experiencia en Java presentan mayor productividad ya que se puede deducir de la gráfica que tanto la mediana como la media son superiores para la categoría Positiva. El grafico de dispersión de la parte derecha de la figura se puede apreciar que tener mayor experiencia en Java no necesariamente implica una mayor productividad. Igual que en los casos anteriores el análisis dicotómico difiere del análisis de correlación de Pearson en cuanto a los resultados obtenidos.



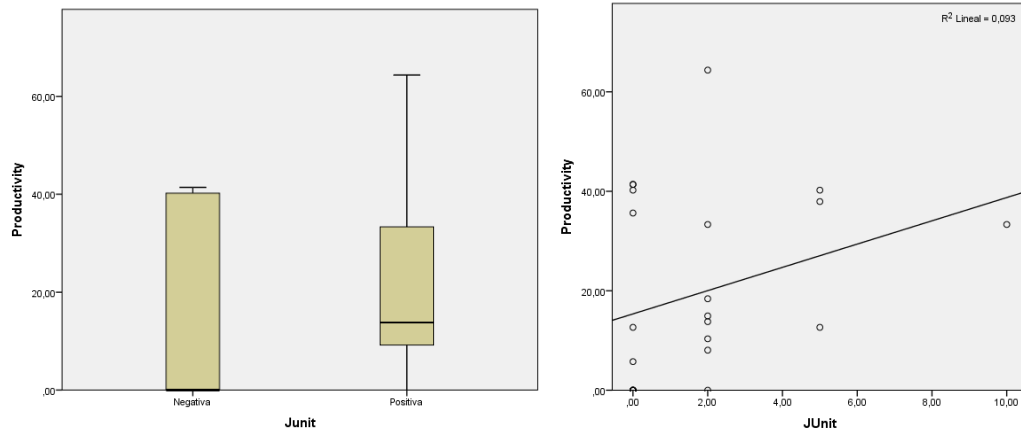
1.4 PRODUCTIVIDAD – EXPERIENCIA EN TESTING

En este caso el análisis dicotómico demuestra que hay una influencia clara de la experiencia en Testing frente sobre la productividad. A simple vista se puede apreciar la diferencia de medias y de las medianas de los dos grupos (Positiva, Negativa). Sin embargo en el análisis de correlación de Pearson se puede ver que no hay indicios para retener esta hipótesis.



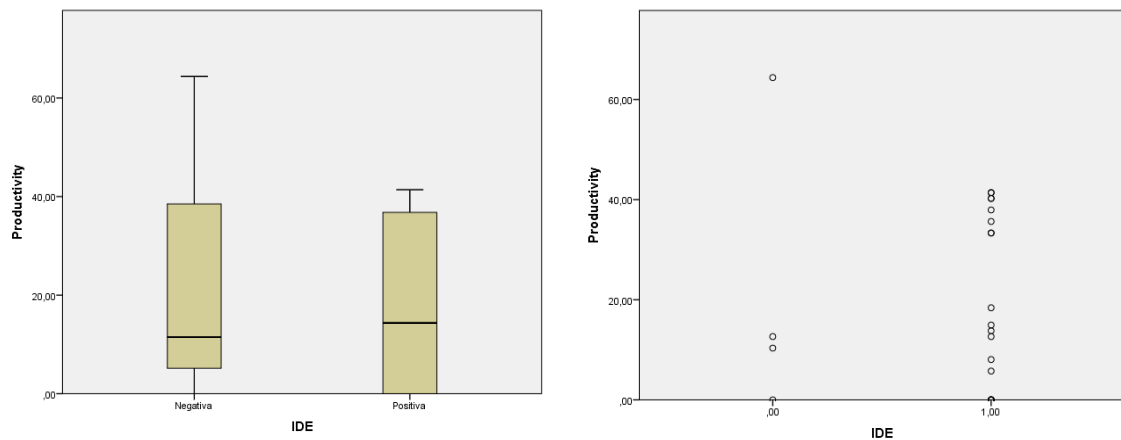
1.5 PRODUCTIVIDAD – EXPERIENCIA EN JUNIT

En el análisis dicotómico se obtiene que la diferencia de medias no es significativa. Tampoco en en análisis de correlación de Pearson, en la gráfica de dispersión se puede apreciar que no hay una relación clara entre la experiencia en JUnit de los sujetos frente a la productividad. Además de que hay dos datos que se pueden considerar dispersos y que condicionan la pendiente de la recta de ajuste lineal.



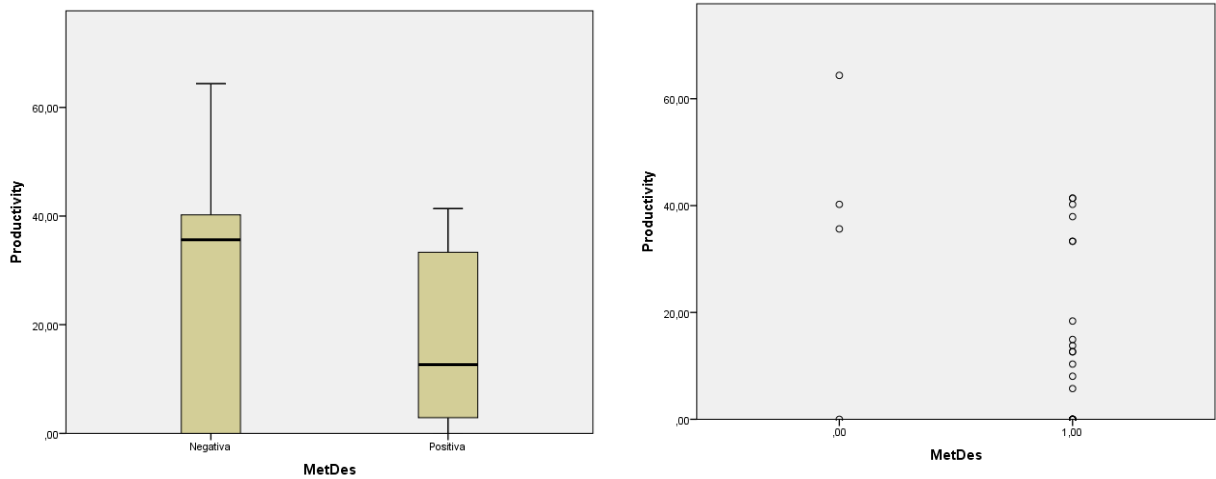
1.6 PRODUCTIVIDAD – EXPERIENCIA IDE

Este es un ejemplo claro de que no hay relación clara entre el par de variables. En la dicotomización el resultado es negativo al igual que en la correlación. En el gráfico de dispersión de los datos cuantitativos se puede apreciar que el que uno de los sujetos que presenta mayor productividad, sin embargo, está claro que si se omite ese dato el resultado sería totalmente opuesto.



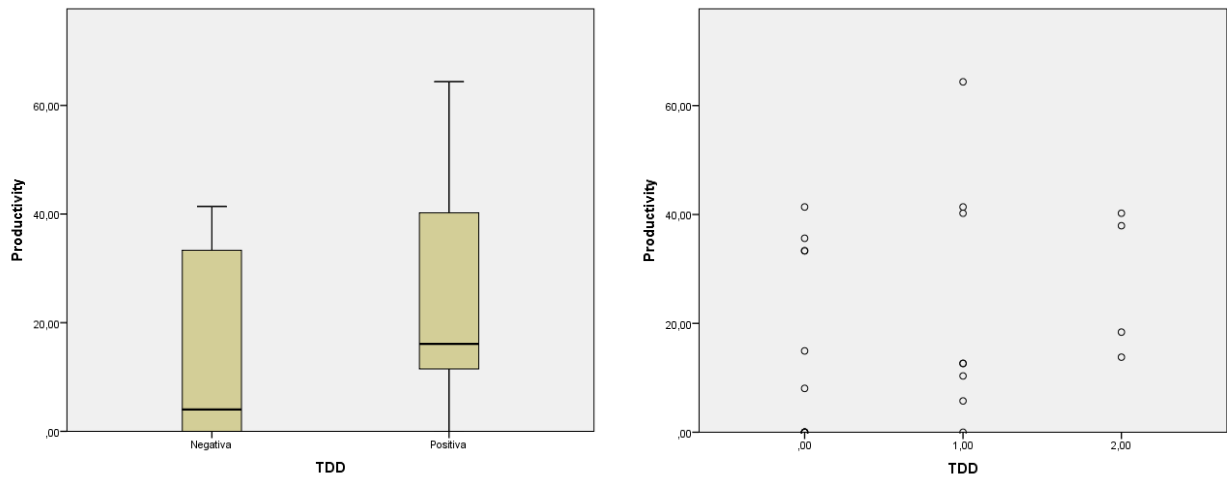
1.7 PRODUCTIVIDAD – EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO

Para la variable metodologías de desarrollo pasa lo mismo que para la experiencia en el uso de IDEs.

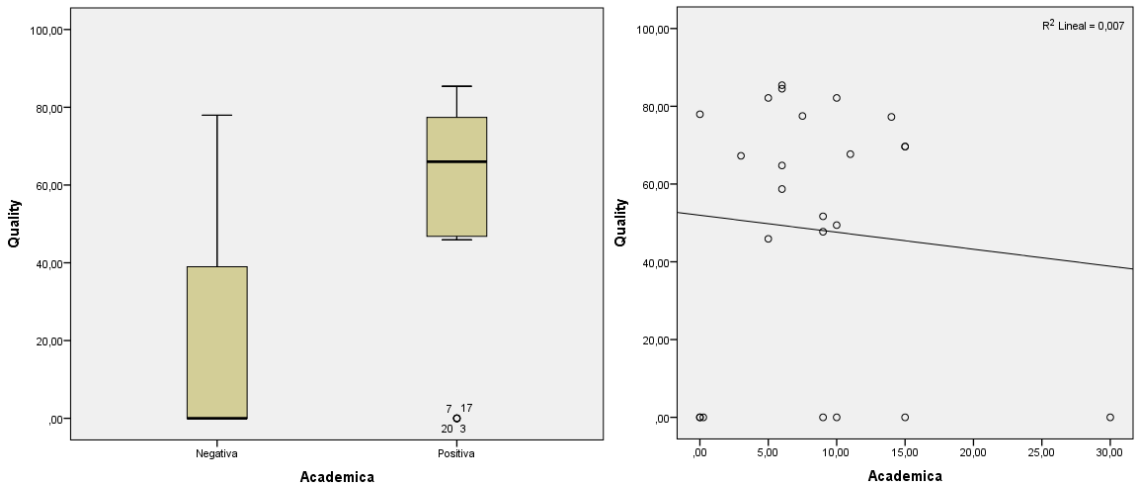


1.8 PRODUCTIVIDAD – EXPERIENCIA EN TDD

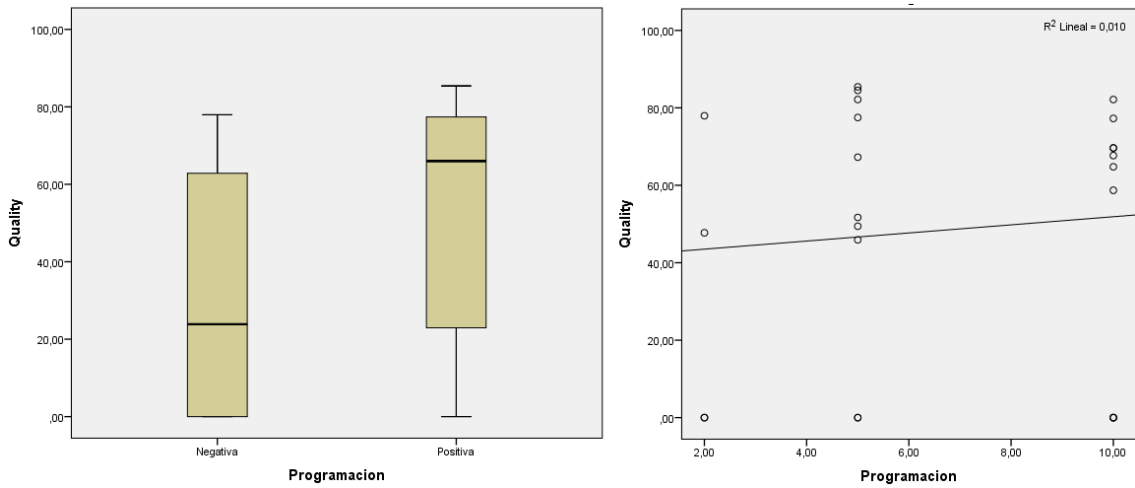
En este caso ni la dicotomización ni la correlación afirman que haya influencia de la experiencia en TDD sobre la productividad.



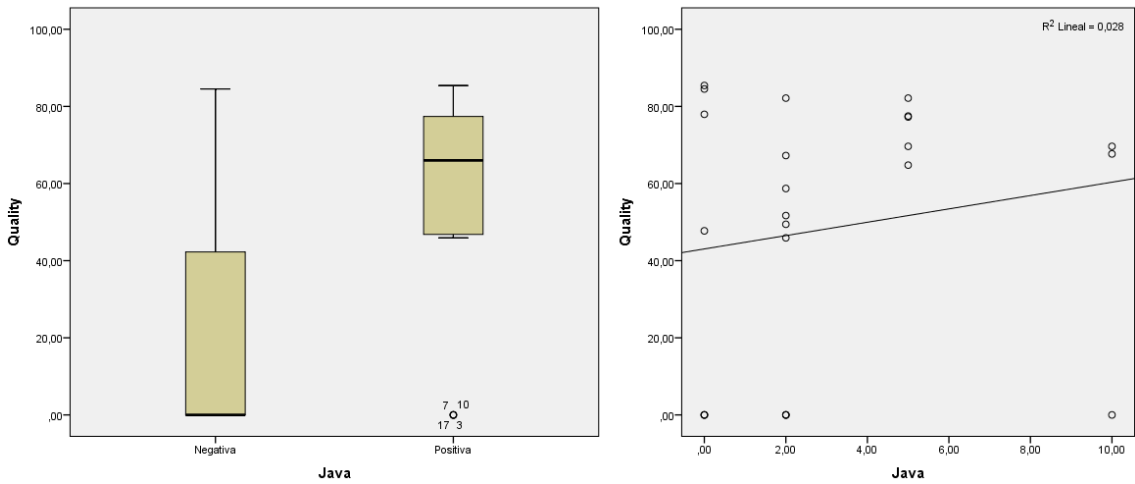
1.9 CALIDAD – EXPERIENCIA ACADÉMICA



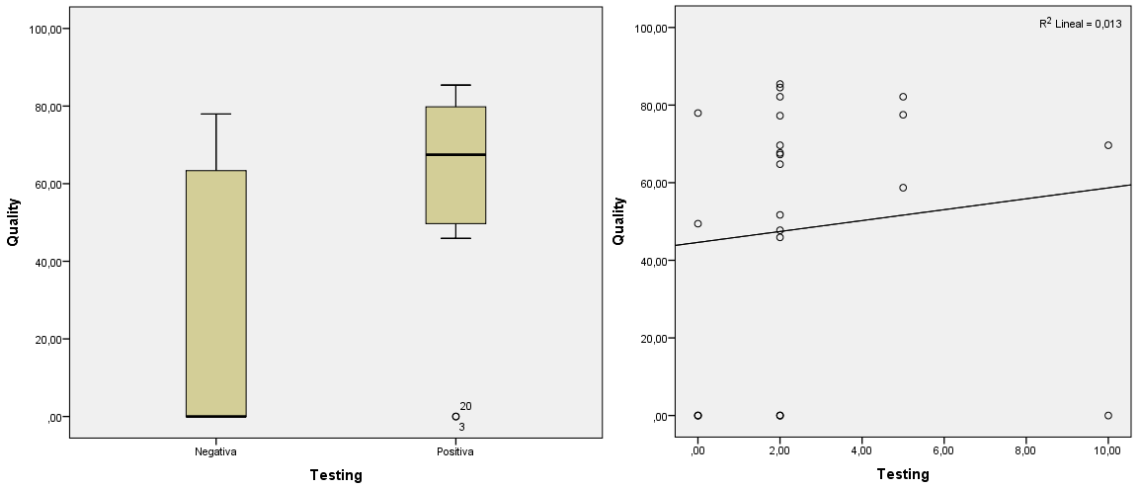
1.10 CALIDAD – EXPERIENCIA DE PROGRAMACIÓN



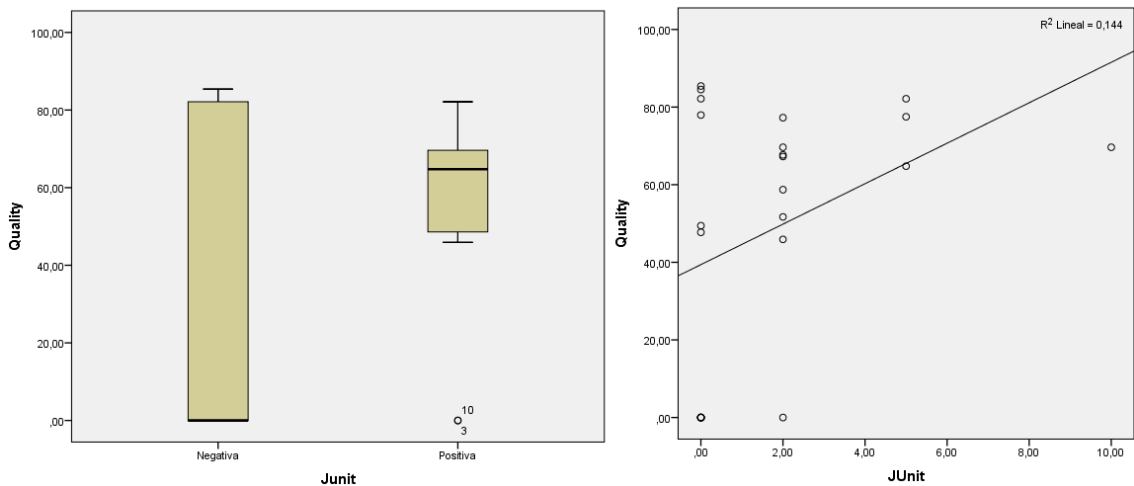
1.11 CALIDAD – EXPERIENCIA EN JAVA



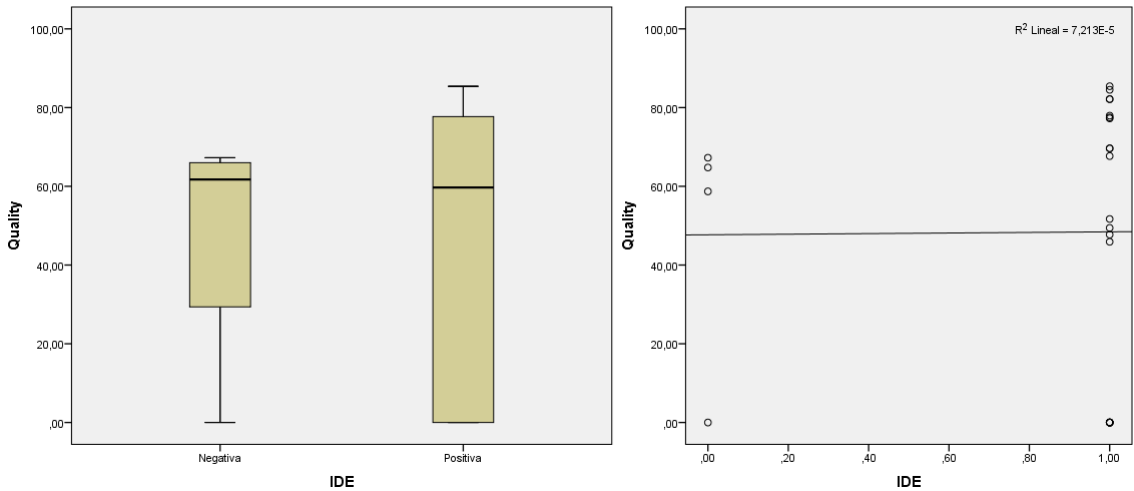
1.12 CALIDAD – EXPERIENCIA EN TESTING



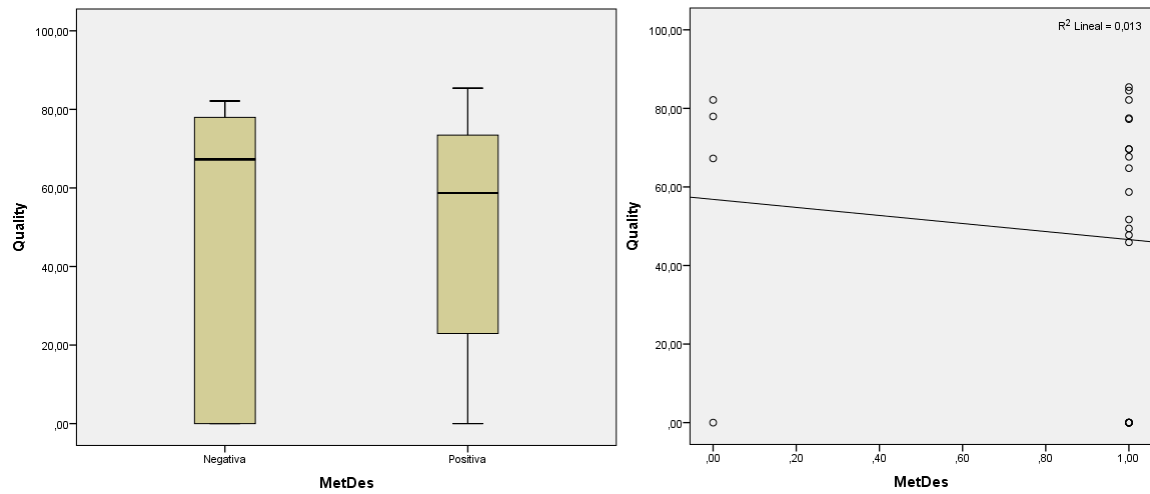
1.13 CALIDAD – EXPERIENCIA EN JUNIT



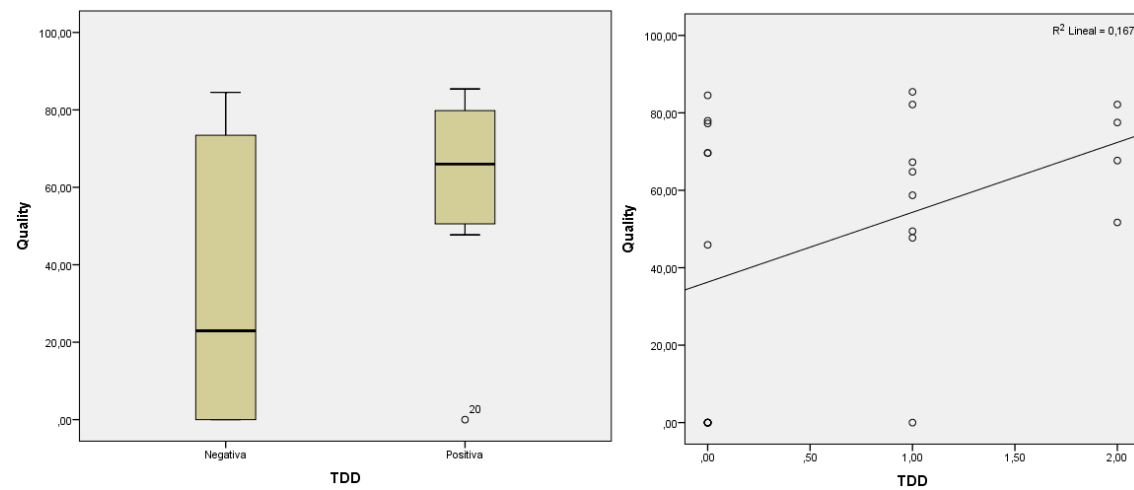
1.14 CALIDAD – EXPERIENCIA IDE



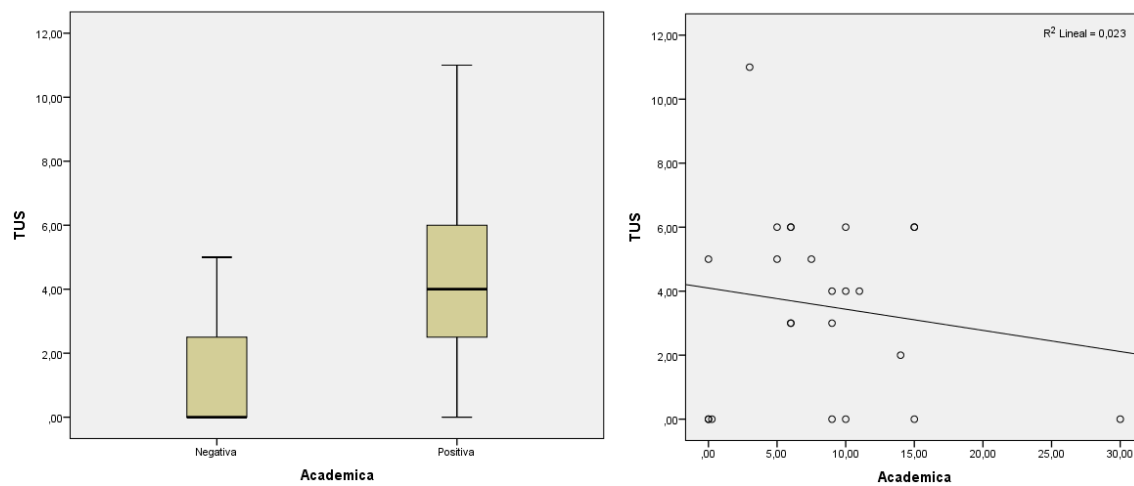
1.15 CALIDAD – EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO



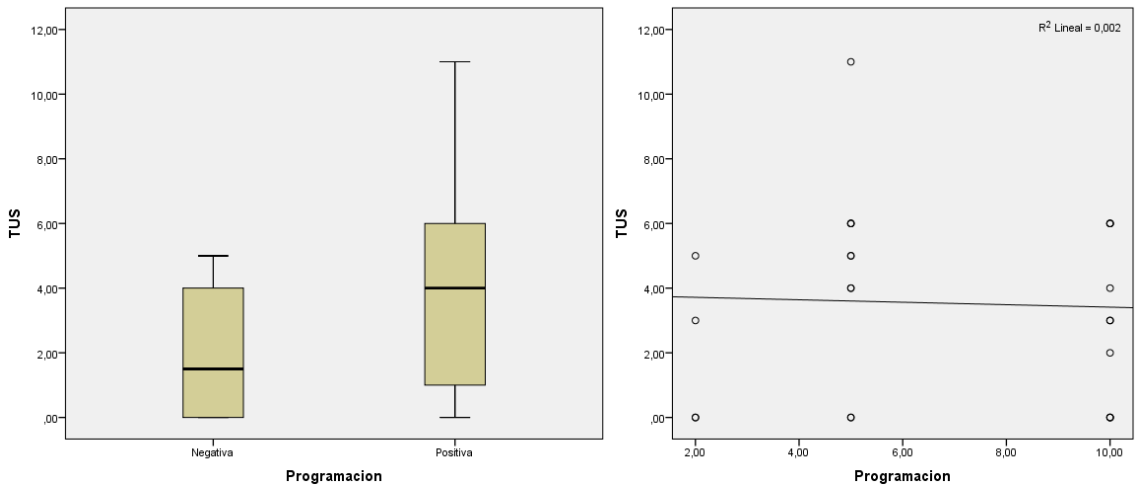
1.16 CALIDAD – EXPERIENCIA EN TDD



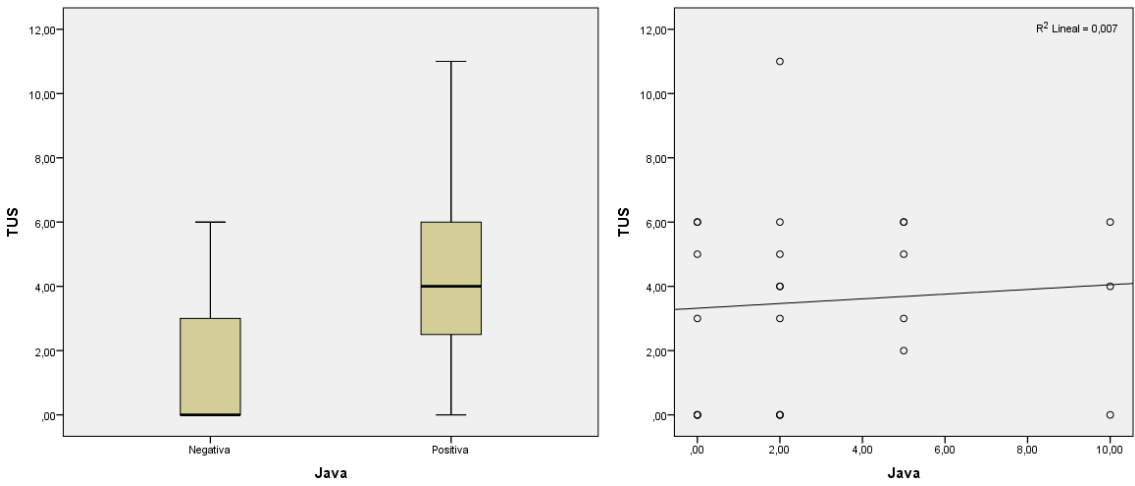
1.17 TUS – EXPERIENCIA ACADÉMICA



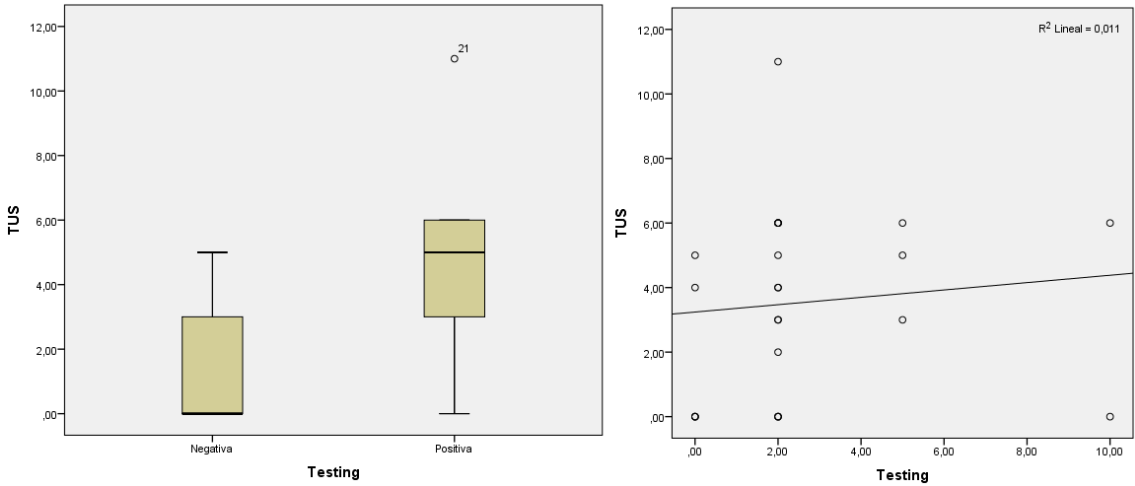
1.18 TUS – EXPERIENCIA DE PROGRAMACIÓN



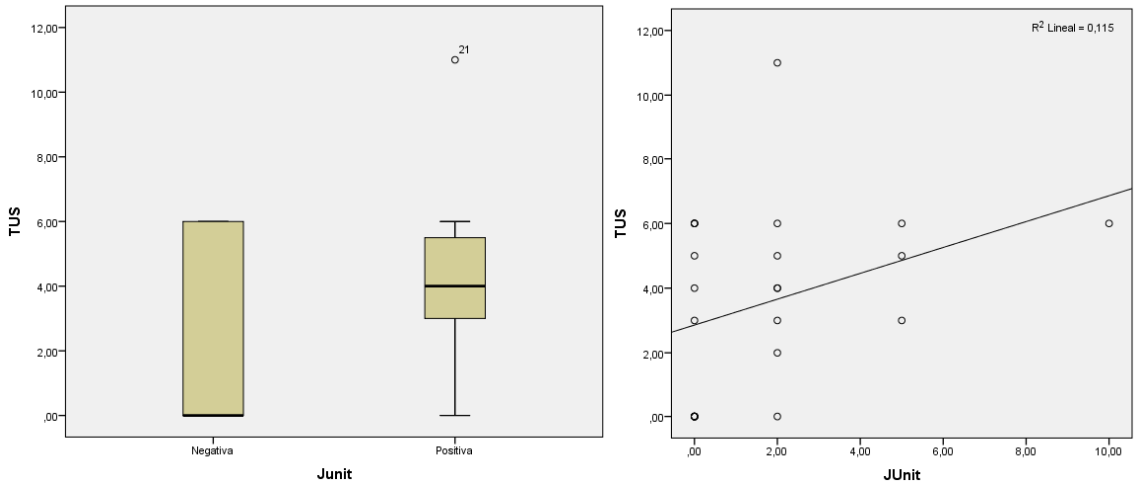
1.19 TUS – EXPERIENCIA EN JAVA



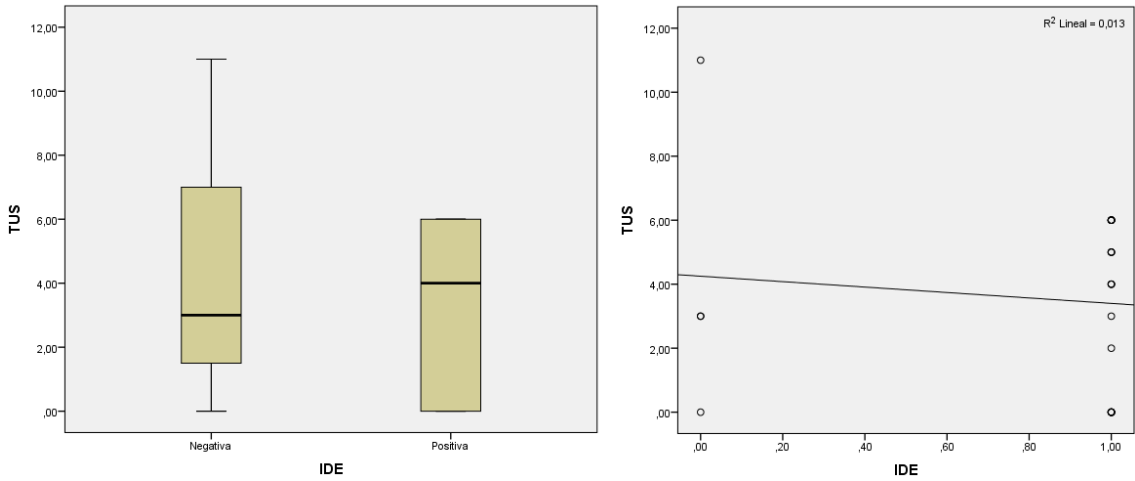
1.20 TUS – EXPERIENCIA EN TESTING



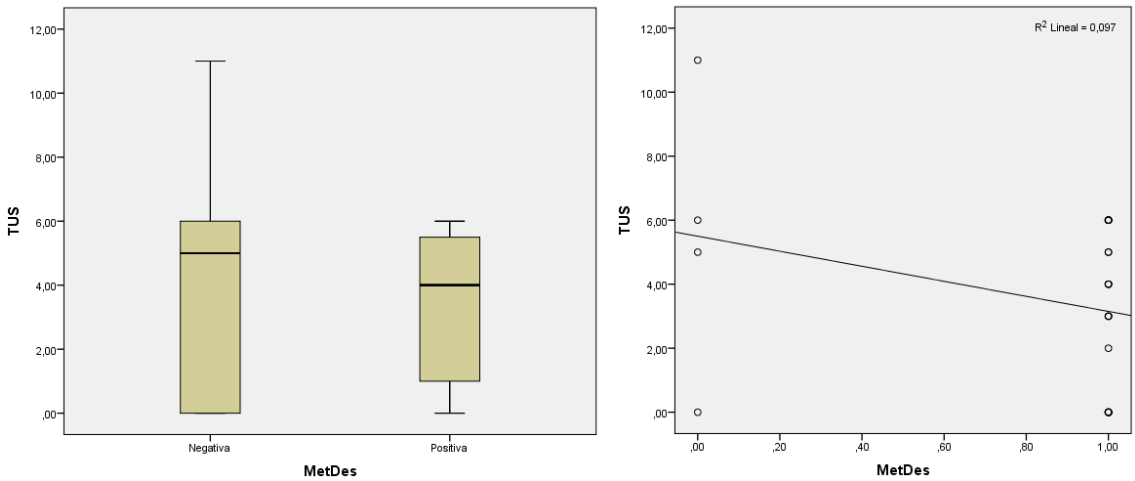
1.21 TUS – EXPERIENCIA EN JUNIT



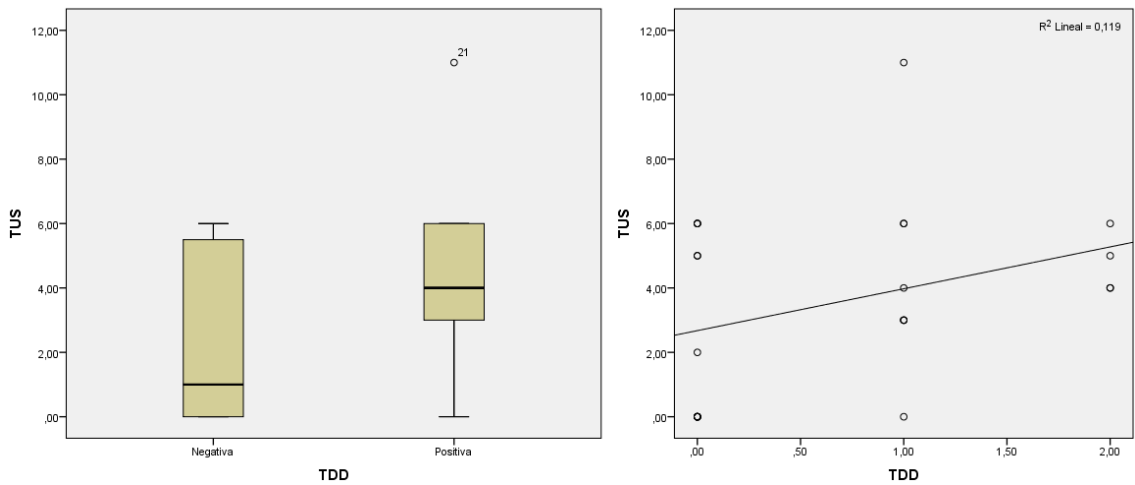
1.22 TUS – EXPERIENCIA IDE



1.23 TUS – EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO

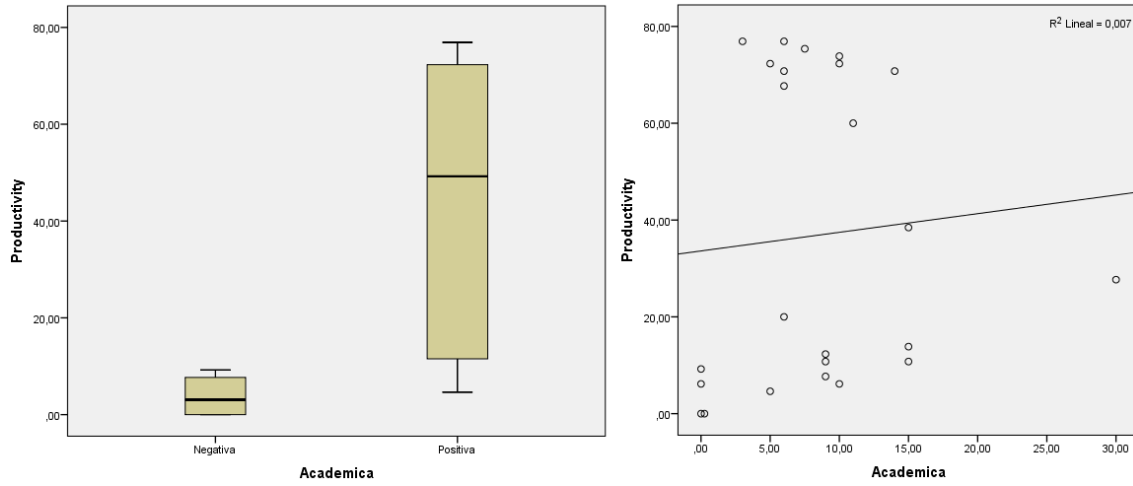


1.24 TUS – EXPERIENCIA EN TDD

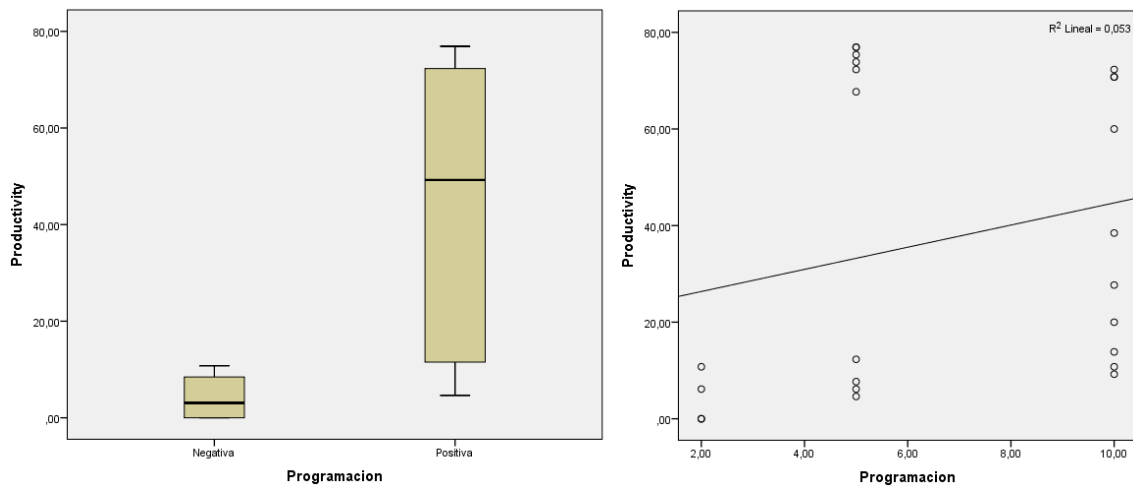


2 DICOTOMIZACIÓN VS CORRELACIÓN BOWLING SCOREKEEPER

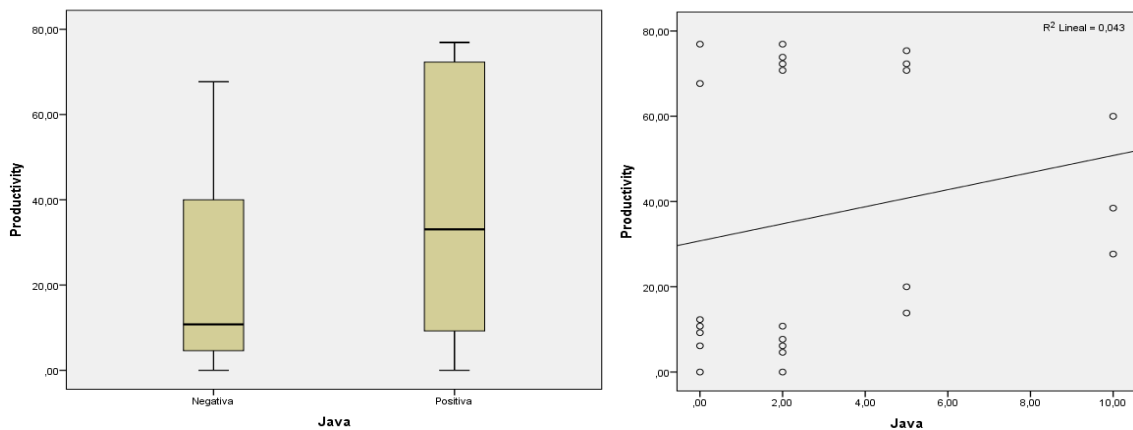
2.1 PRODUCTIVIDAD – EXPERIENCIA ACADÉMICA



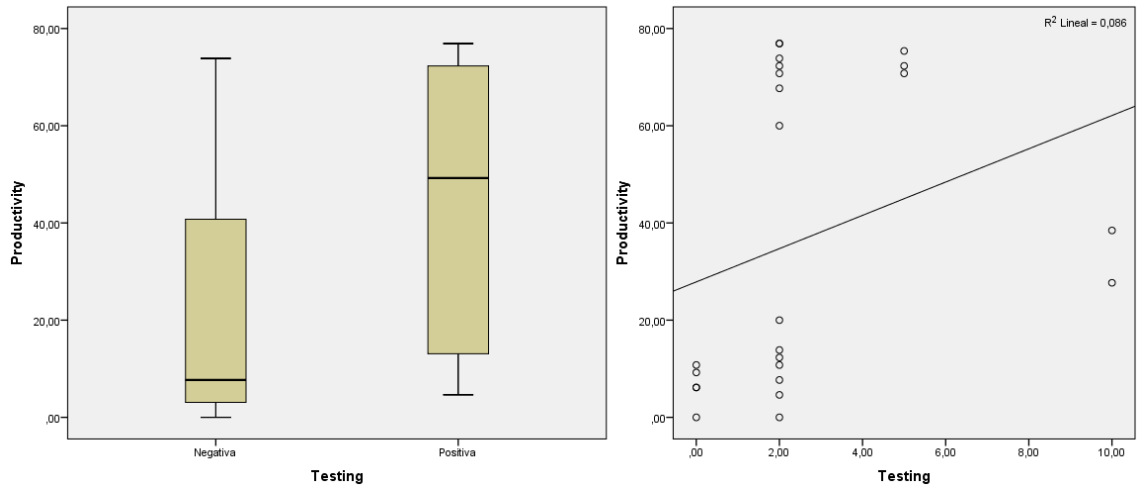
2.2 PRODUCTIVIDAD – EXPERIENCIA DE PROGRAMACIÓN



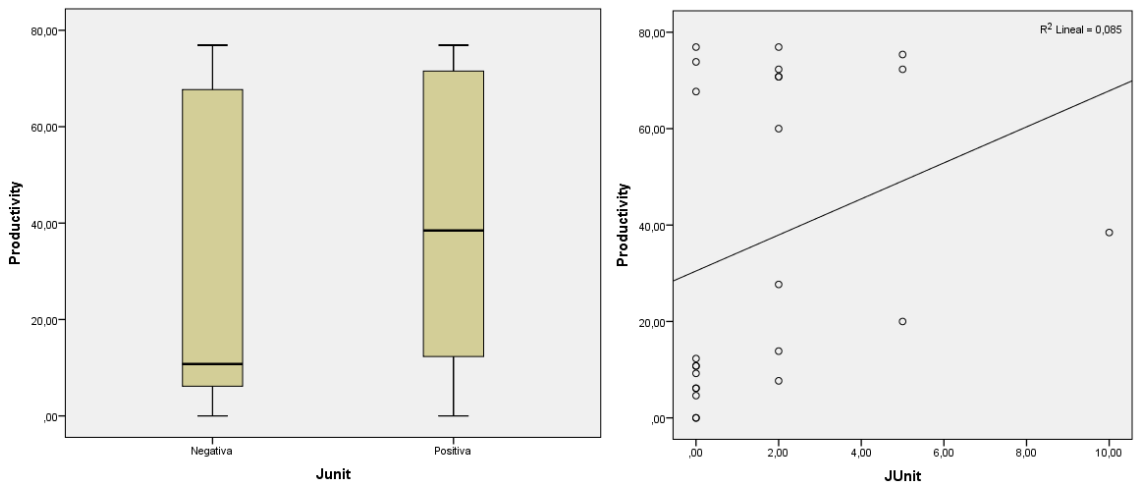
2.3 PRODUCTIVIDAD – EXPERIENCIA EN JAVA



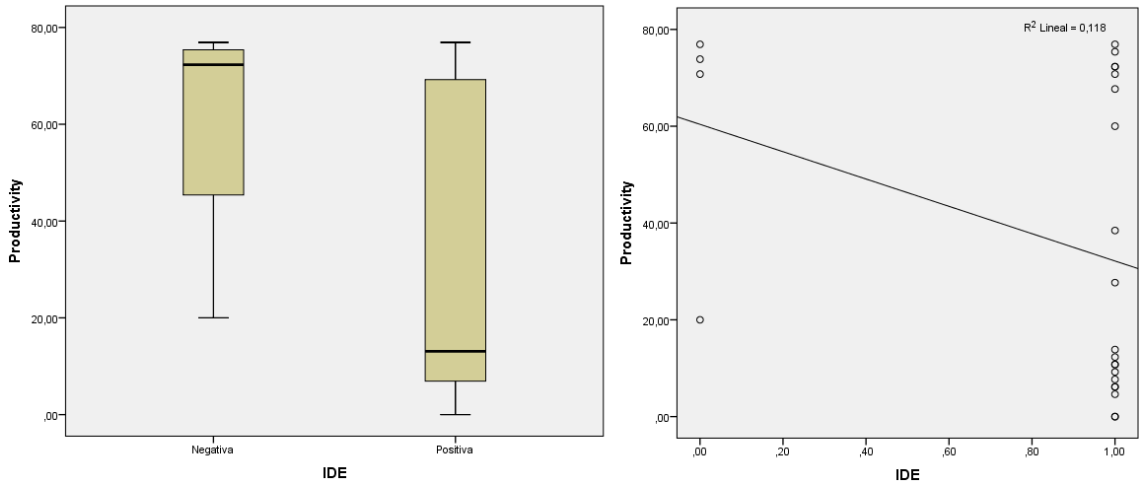
2.4 PRODUCTIVIDAD – EXPERIENCIA EN TESTING



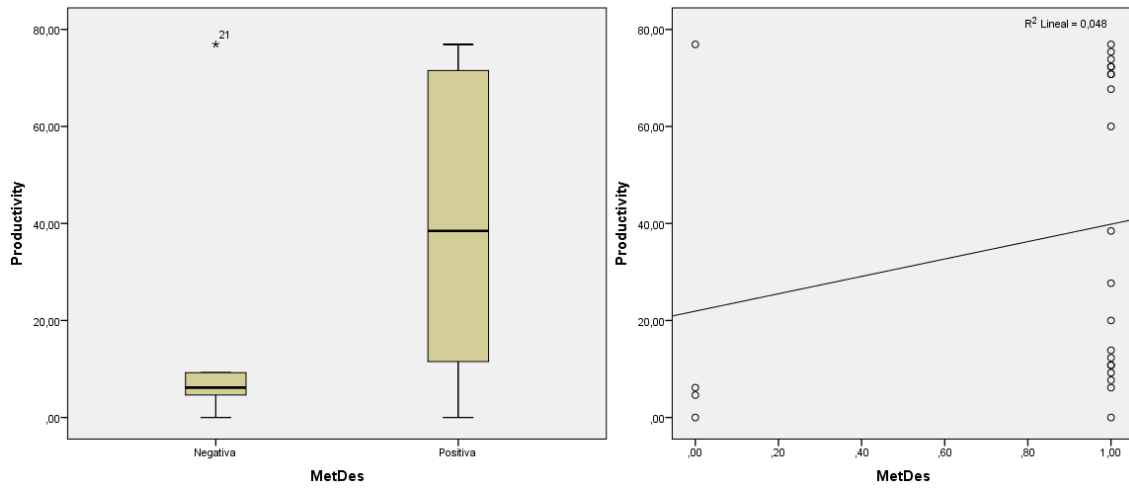
2.5 PRODUCTIVIDAD – EXPERIENCIA EN JUNIT



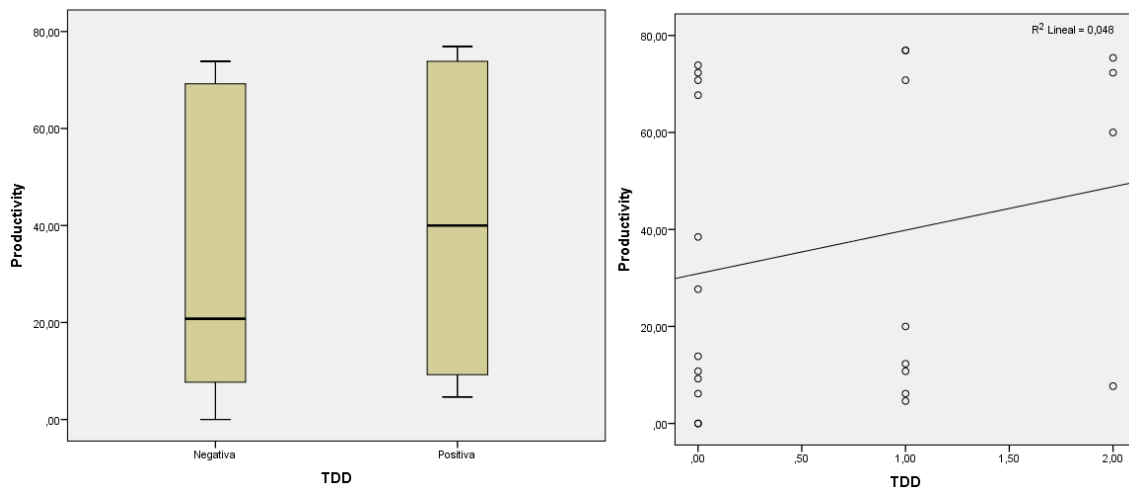
2.6 PRODUCTIVIDAD – EXPERIENCIA IDE



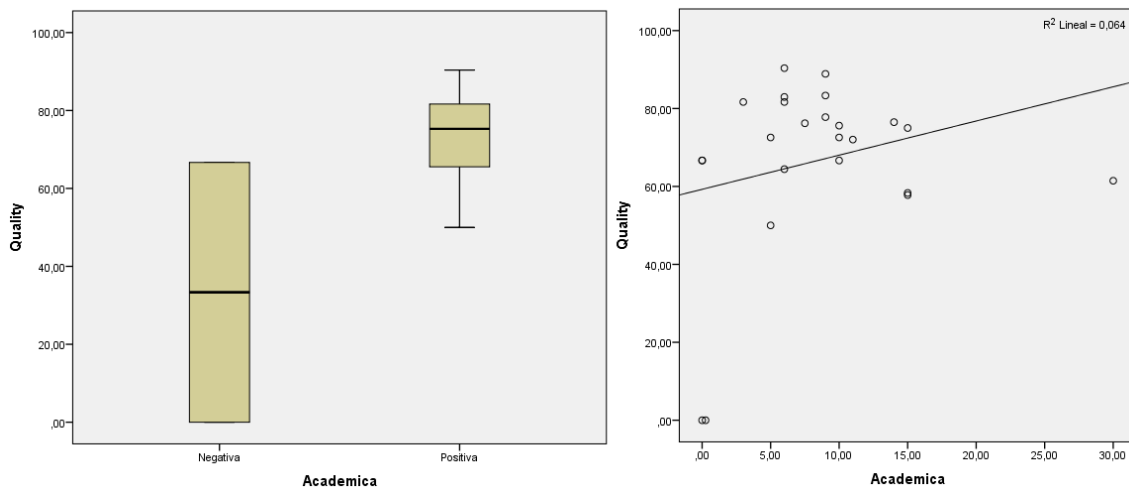
2.7 PRODUCTIVIDAD – EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO



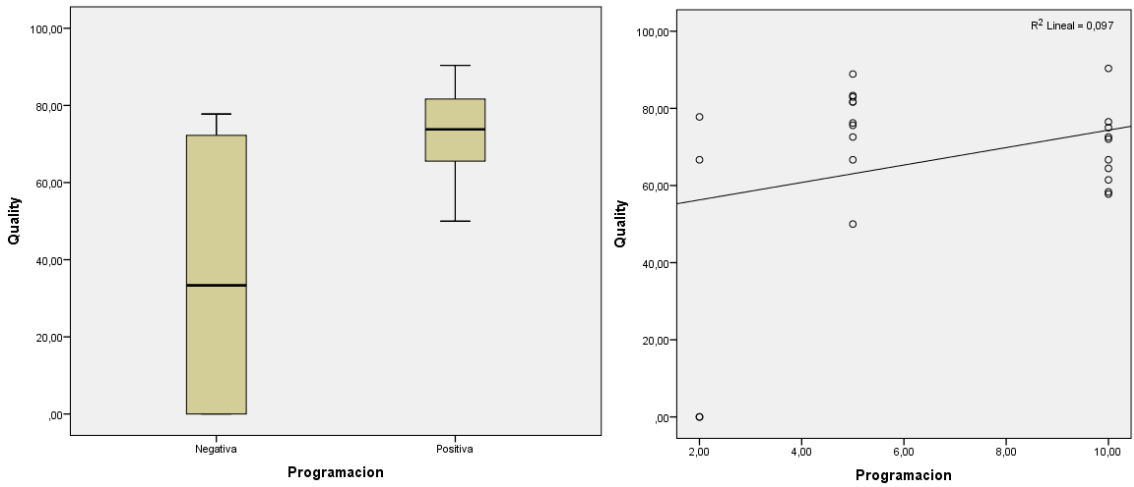
2.8 PRODUCTIVIDAD – EXPERIENCIA EN TDD



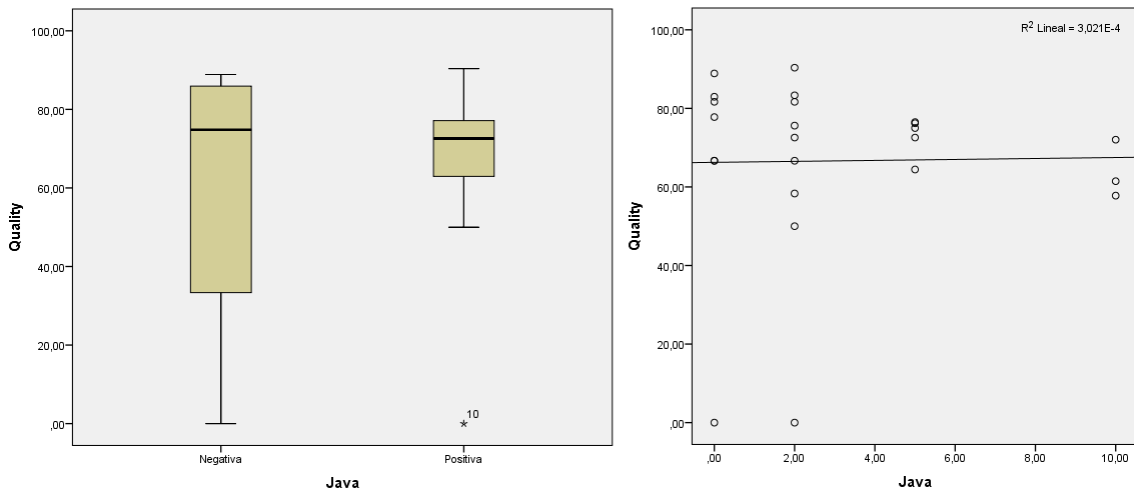
2.9 CALIDAD – EXPERIENCIA ACADÉMICA



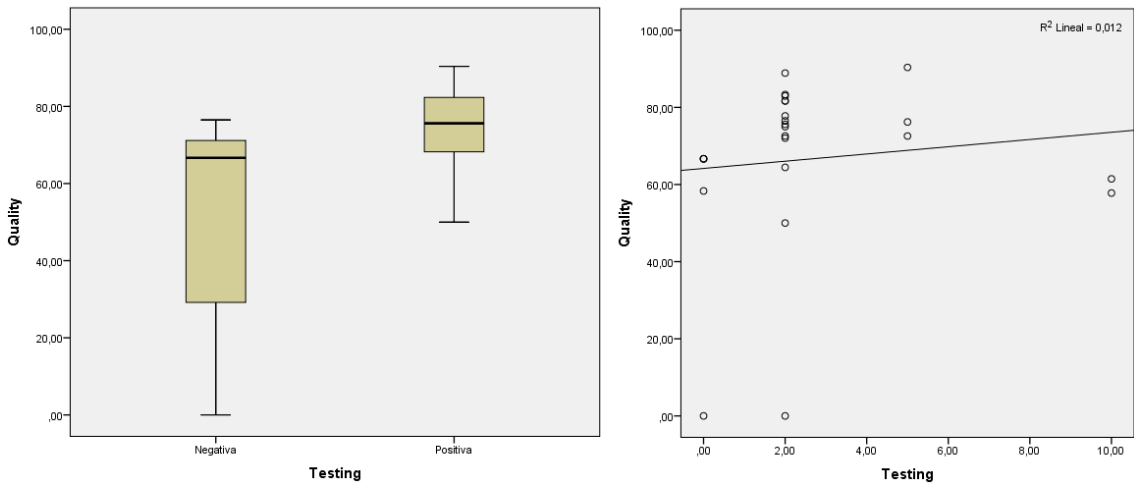
2.10 CALIDAD – EXPERIENCIA DE PROGRAMACIÓN



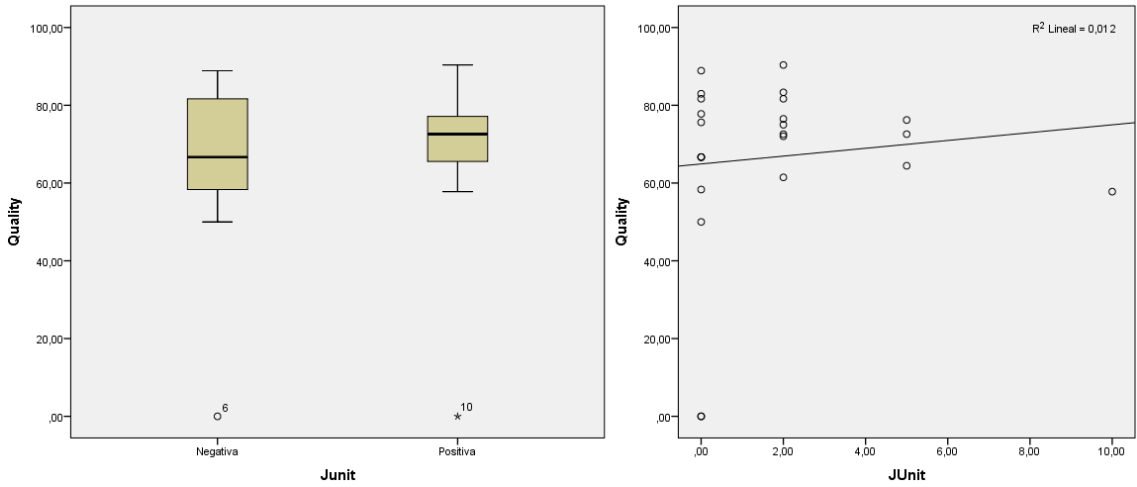
2.11 CALIDAD – EXPERIENCIA EN JAVA



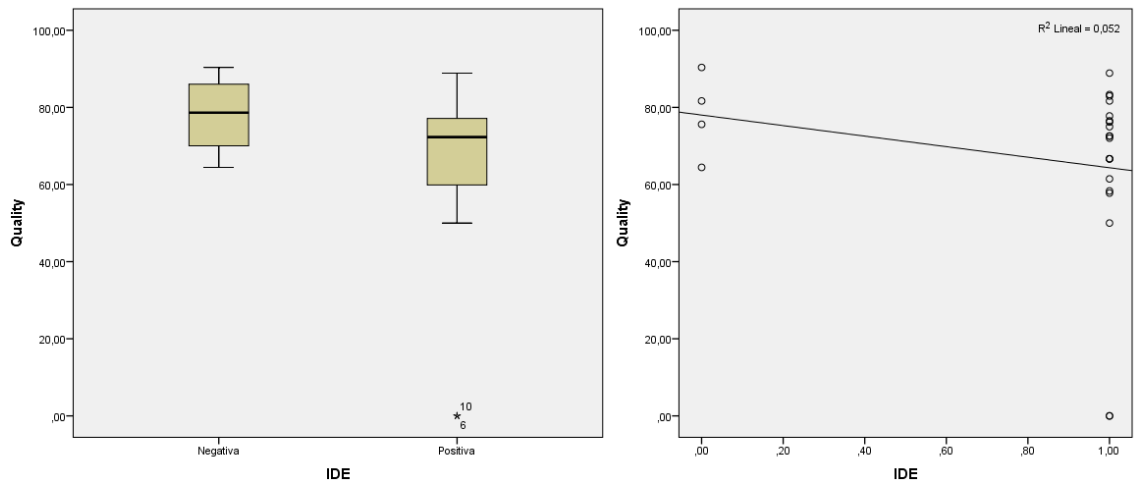
2.12 CALIDAD – EXPERIENCIA EN TESTING



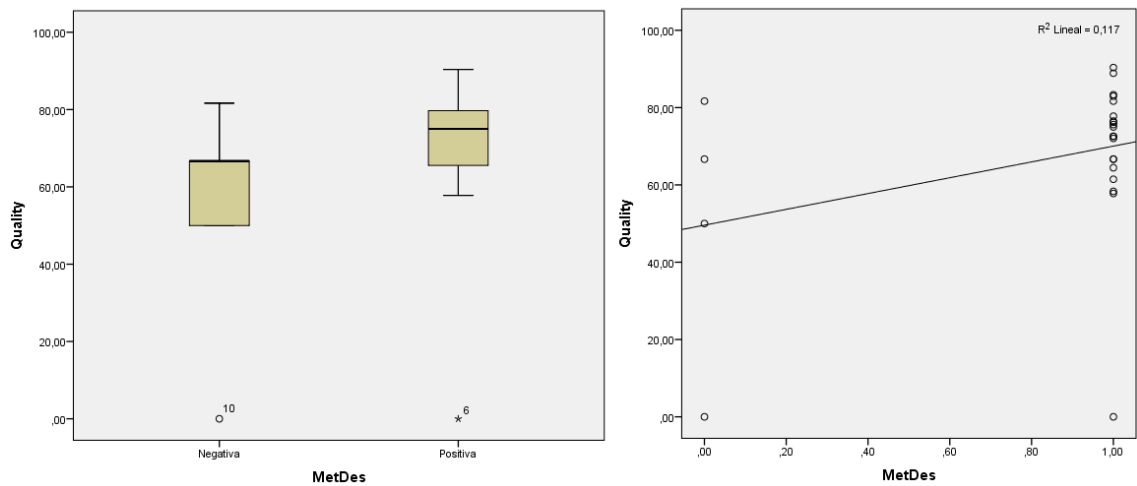
2.13 CALIDAD – EXPERIENCIA EN JUNIT



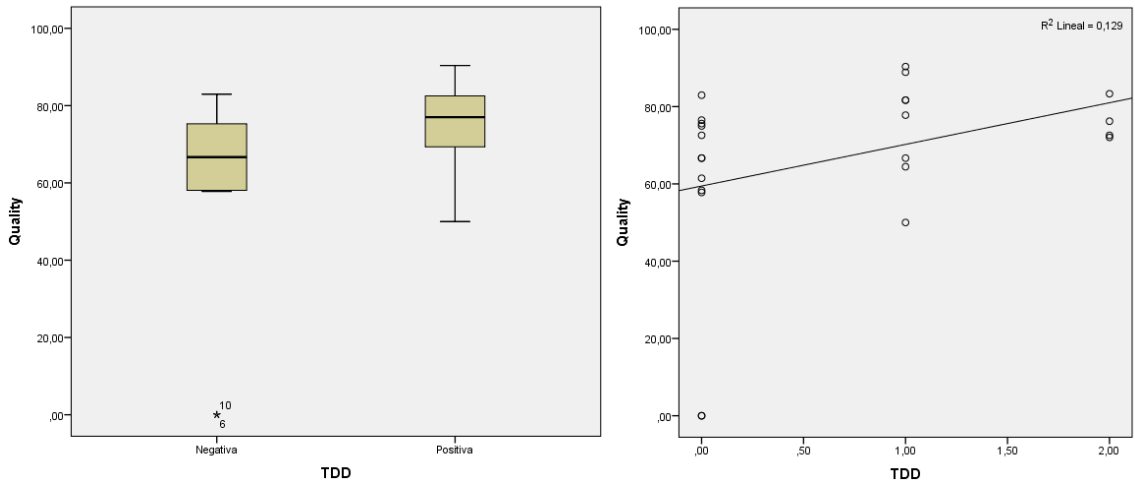
2.14 CALIDAD – EXPERIENCIA IDE



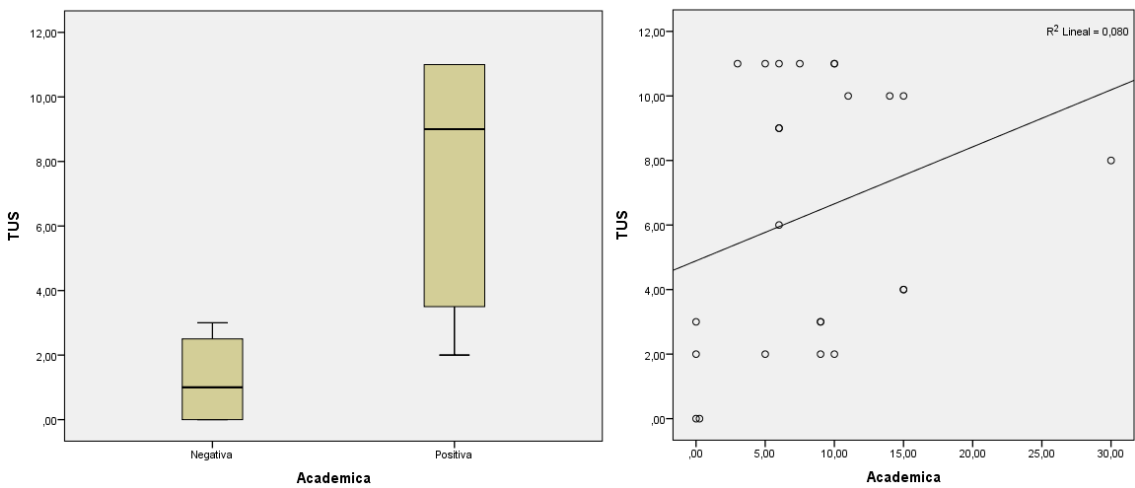
2.15 CALIDAD – EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO



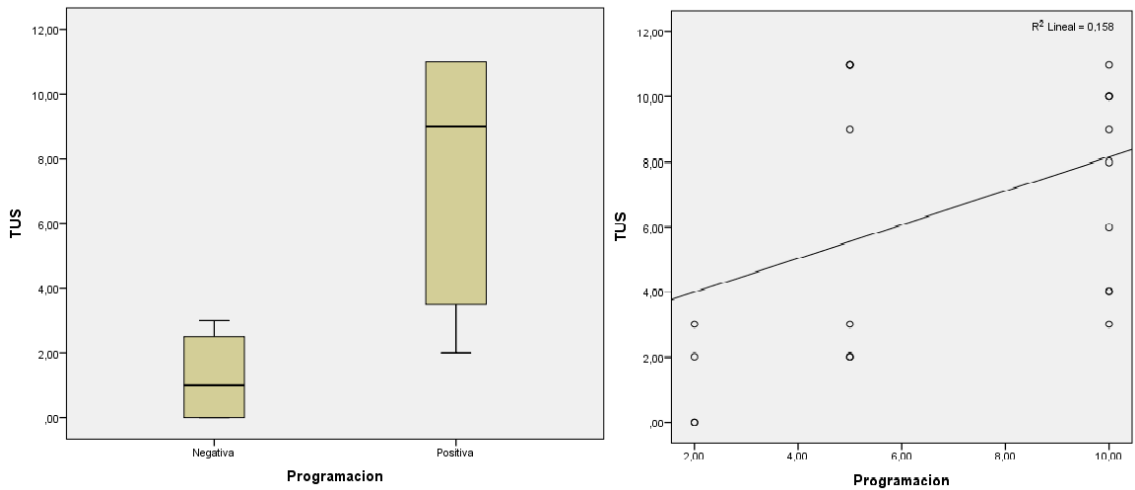
2.16 CALIDAD – EXPERIENCIA EN TDD



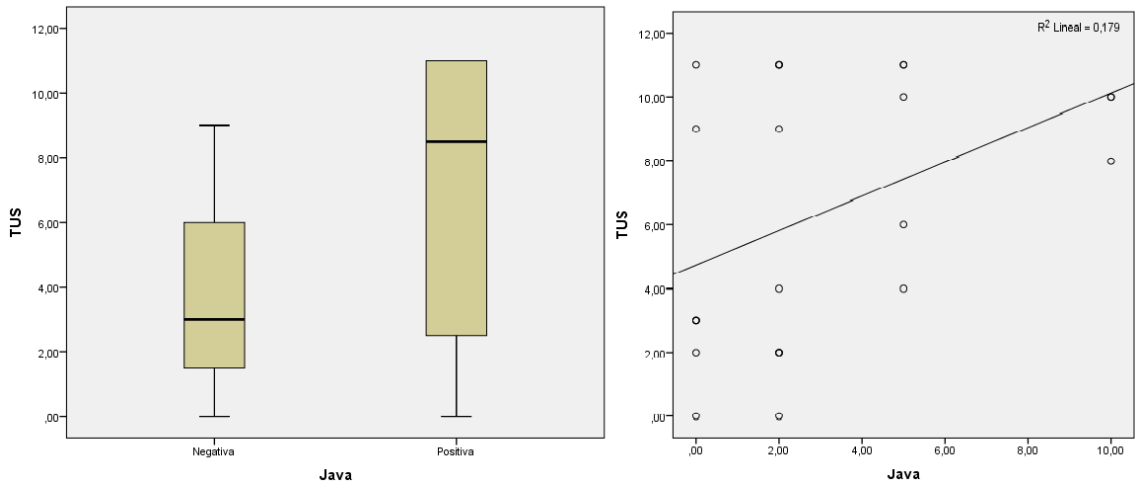
2.17 TUS – EXPERIENCIA ACADÉMICA



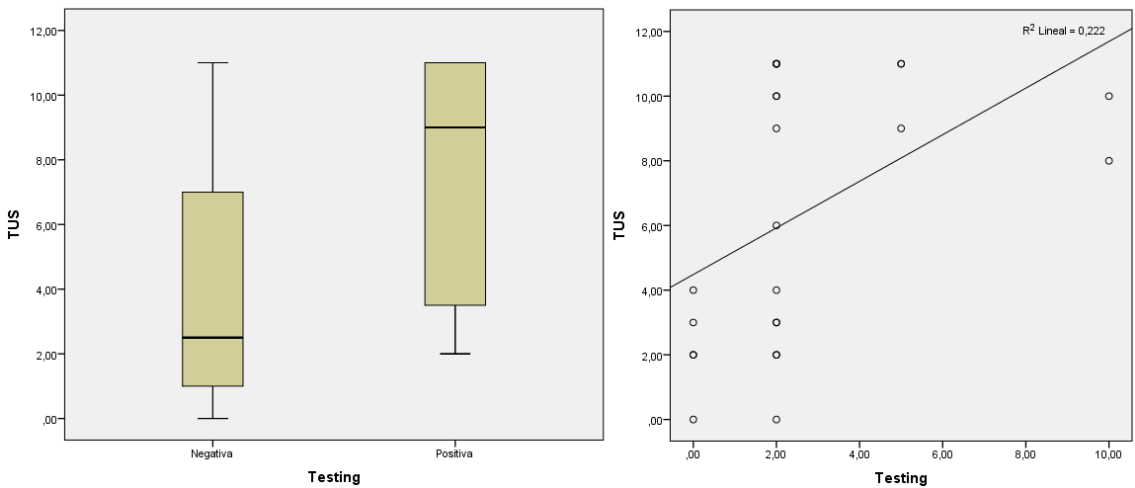
2.18 TUS – EXPERIENCIA DE PROGRAMACIÓN



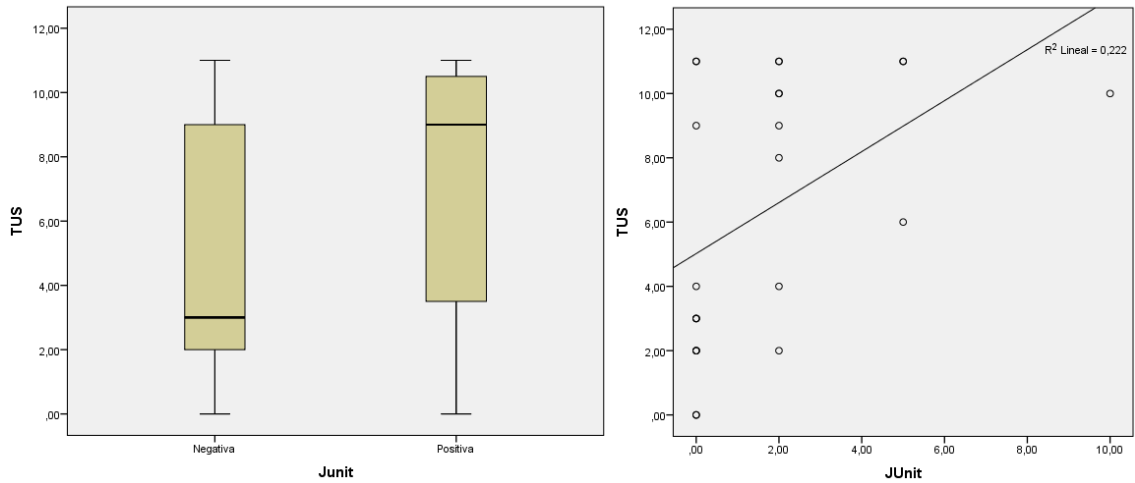
2.19 TUS – EXPERIENCIA EN JAVA



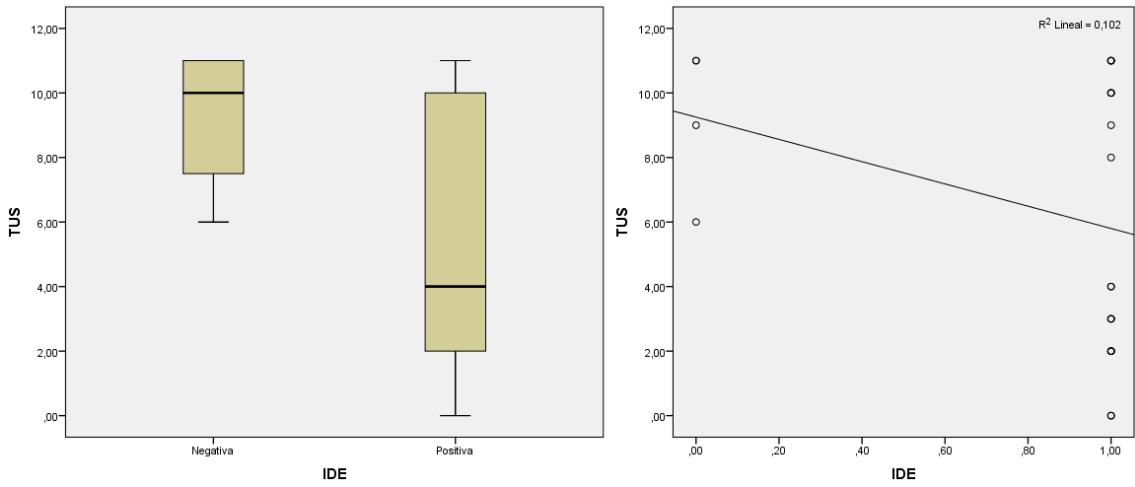
2.20 TUS – EXPERIENCIA EN TESTING



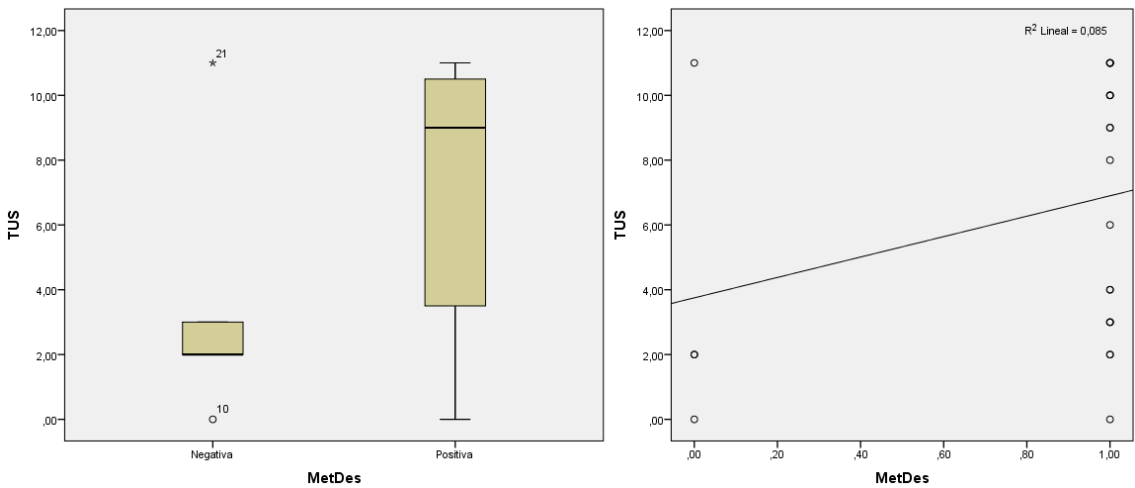
2.21 TUS – EXPERIENCIA EN JUNIT



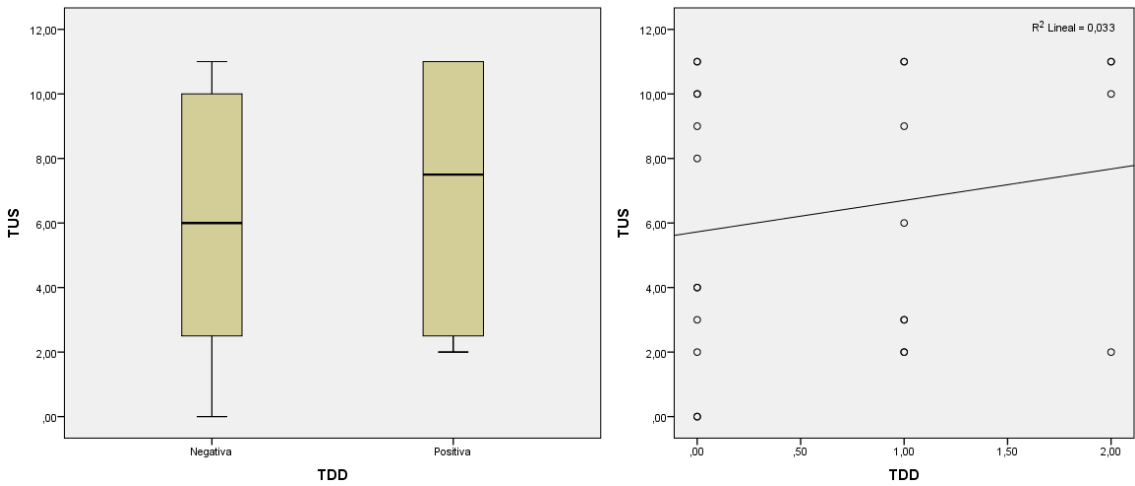
2.22 TUS – EXPERIENCIA IDE



2.23 TUS – EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO

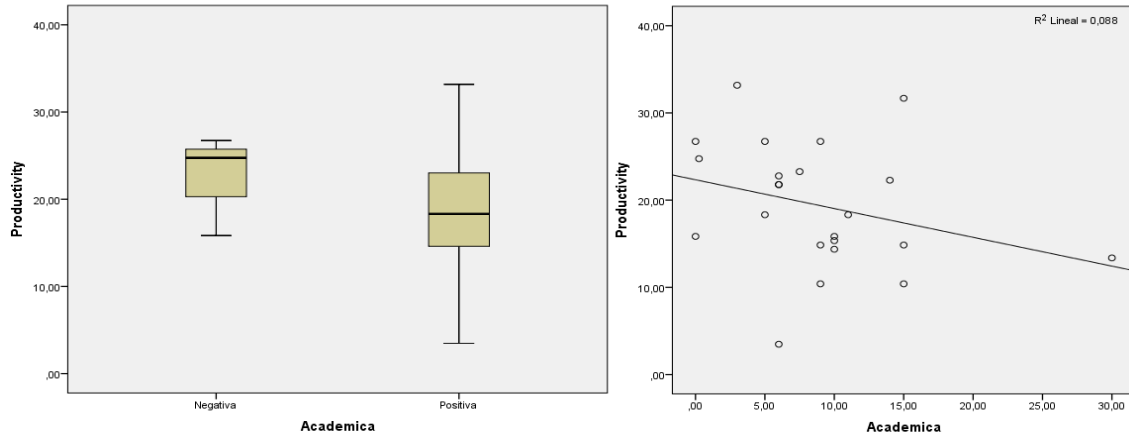


2.24 TUS – EXPERIENCIA EN TDD

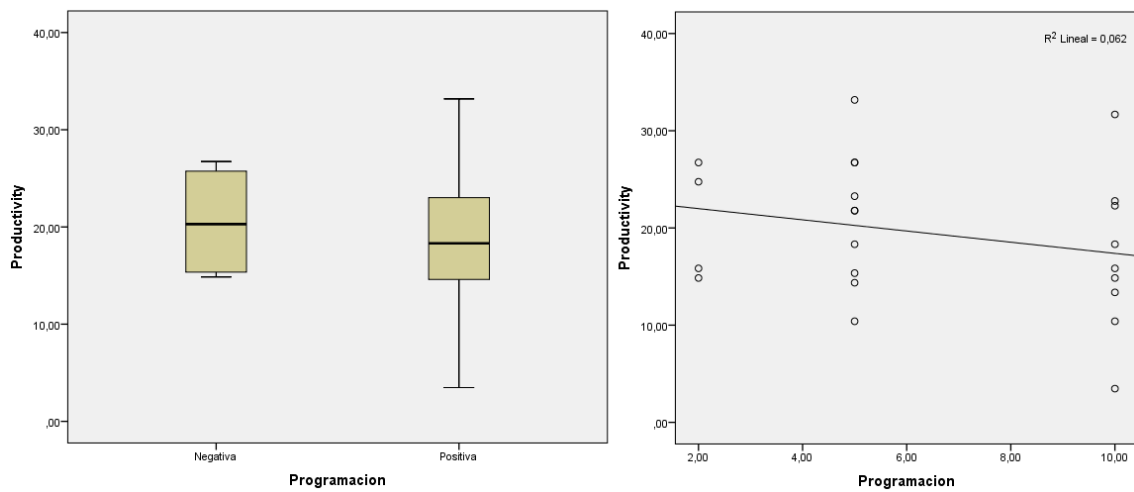


3 DICOTOMIZACIÓN VS CORRELACIÓN MUSICPHONE

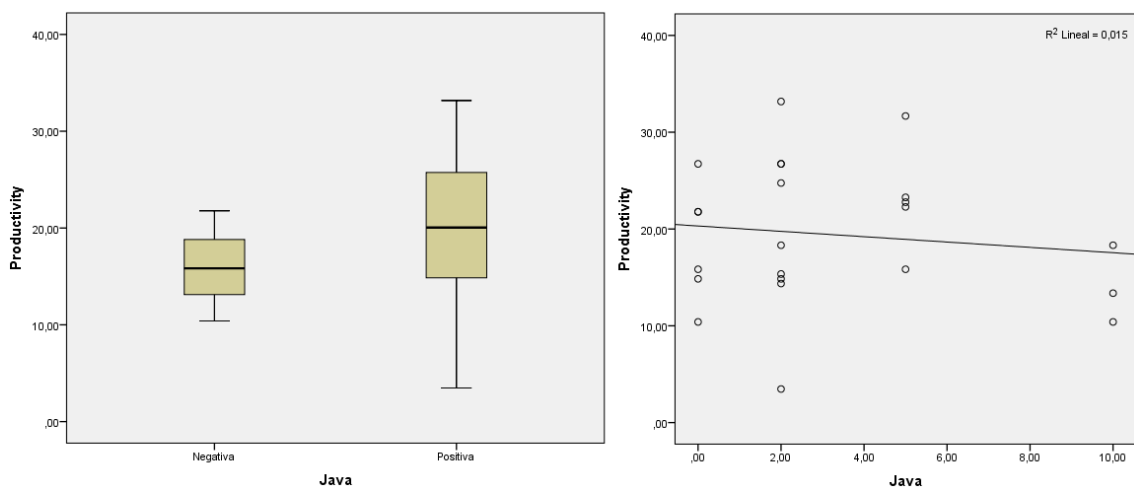
3.1 PRODUCTIVIDAD – EXPERIENCIA ACADÉMICA



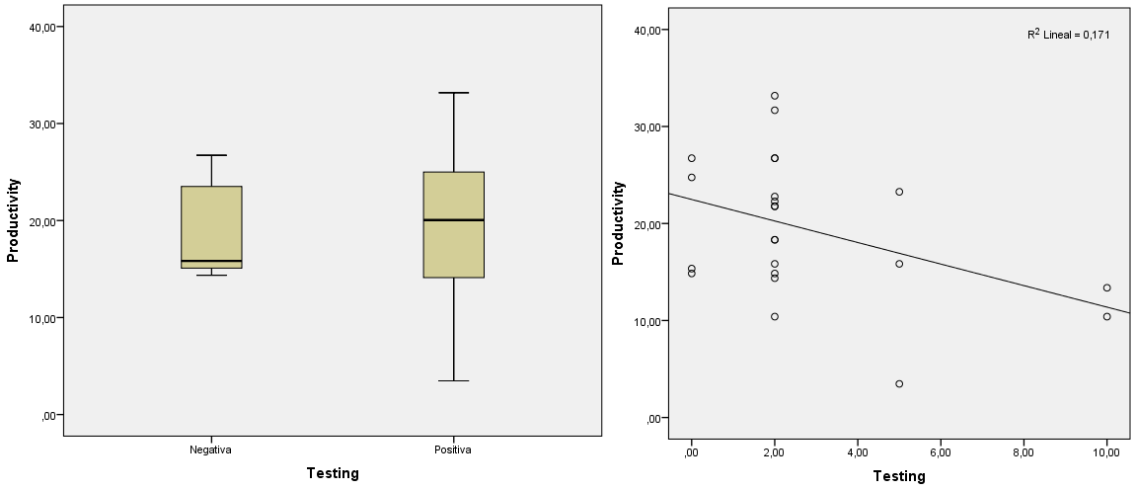
3.2 PRODUCTIVIDAD – EXPERIENCIA DE PROGRAMACIÓN



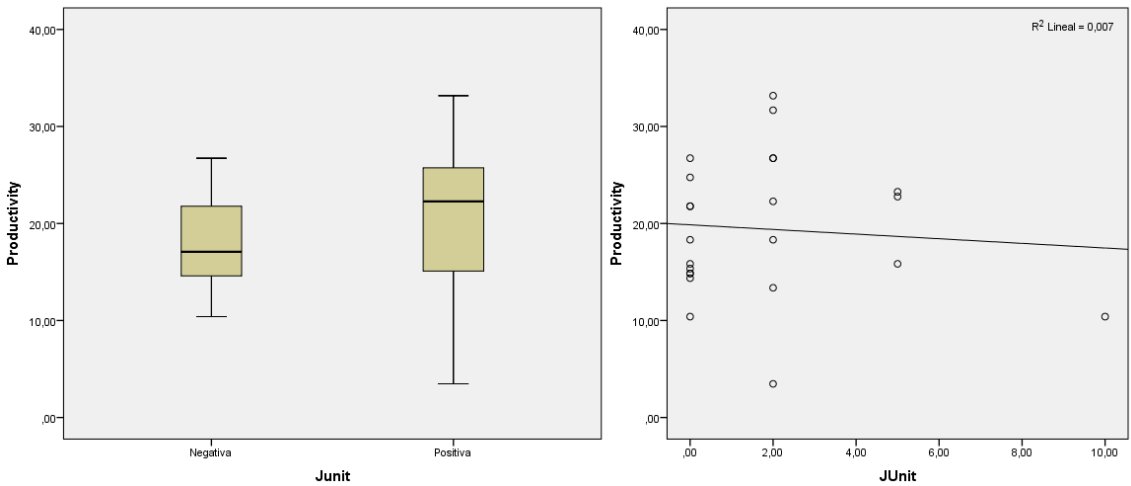
3.3 PRODUCTIVIDAD – EXPERIENCIA EN JAVA



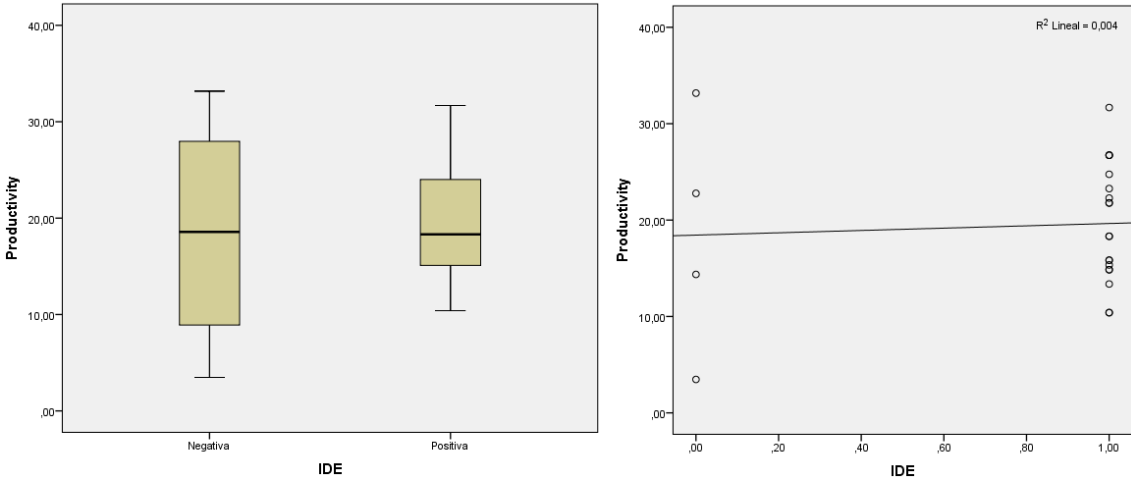
3.4 PRODUCTIVIDAD – EXPERIENCIA EN TESTING



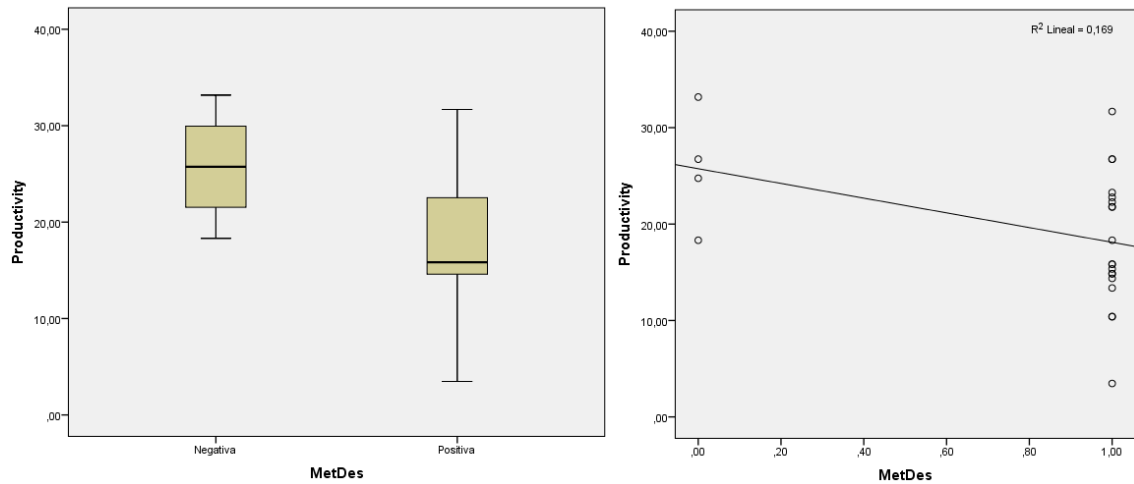
3.5 PRODUCTIVIDAD – EXPERIENCIA EN JUNIT



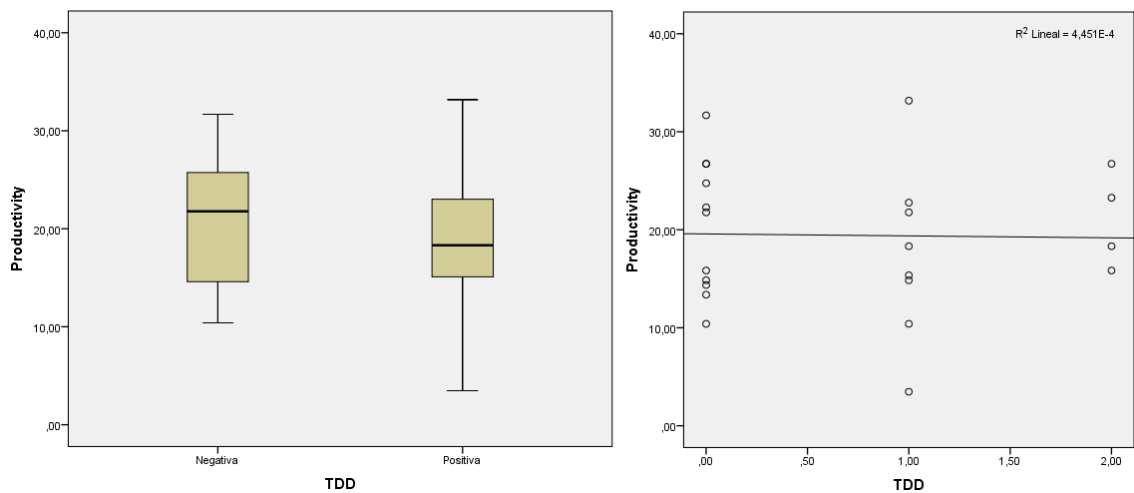
3.6 PRODUCTIVIDAD – EXPERIENCIA IDE



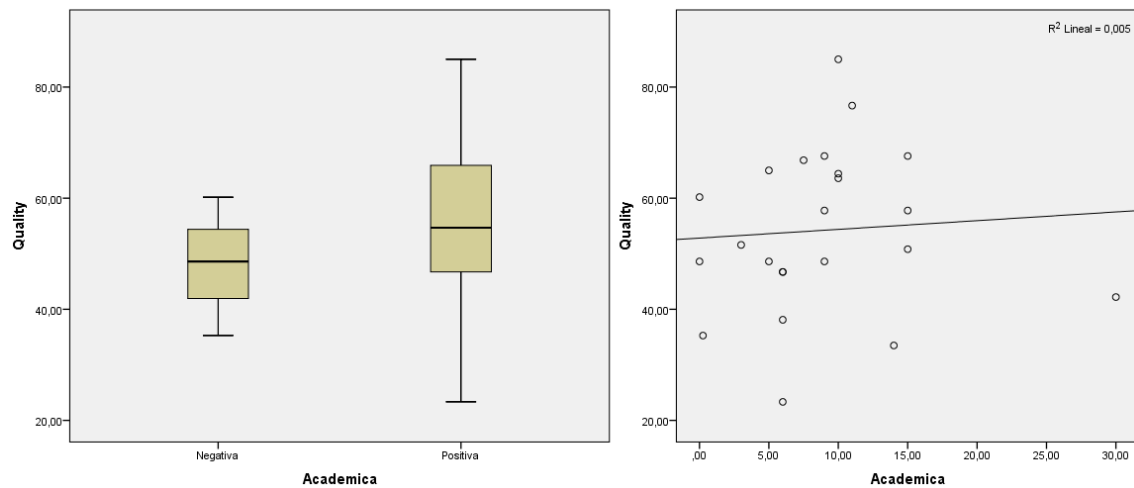
3.7 PRODUCTIVIDAD – EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO



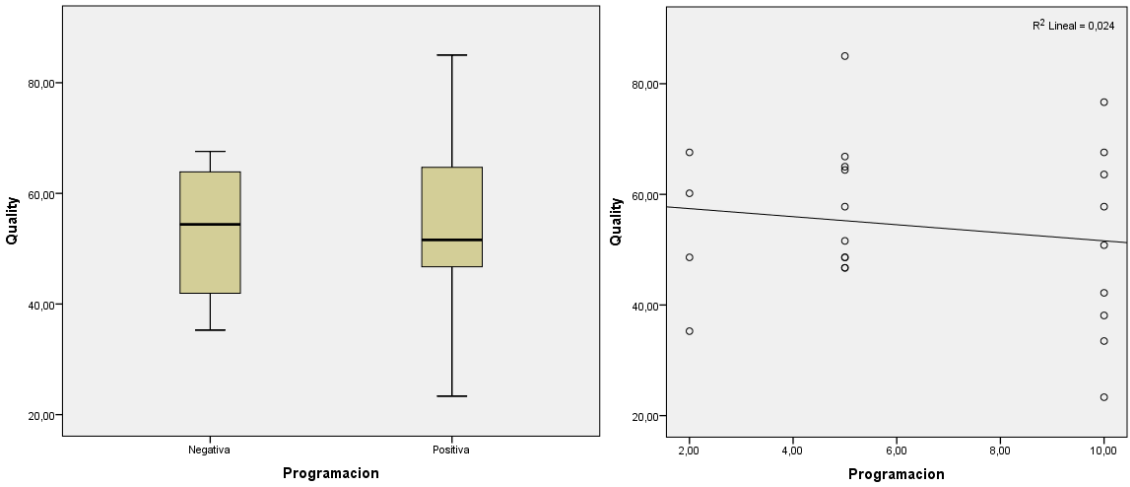
3.8 PRODUCTIVIDAD – EXPERIENCIA EN TDD



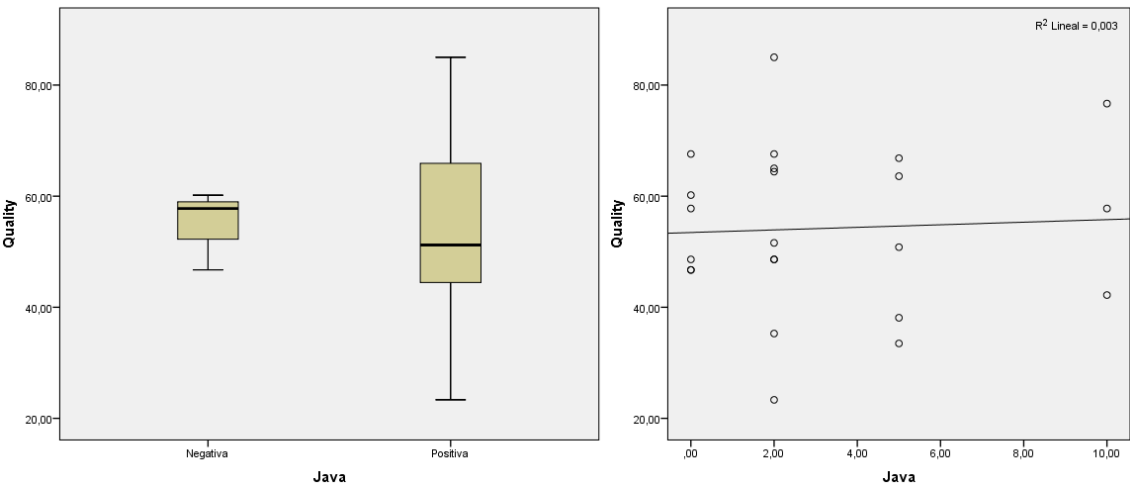
3.9 CALIDAD – EXPERIENCIA ACADÉMICA



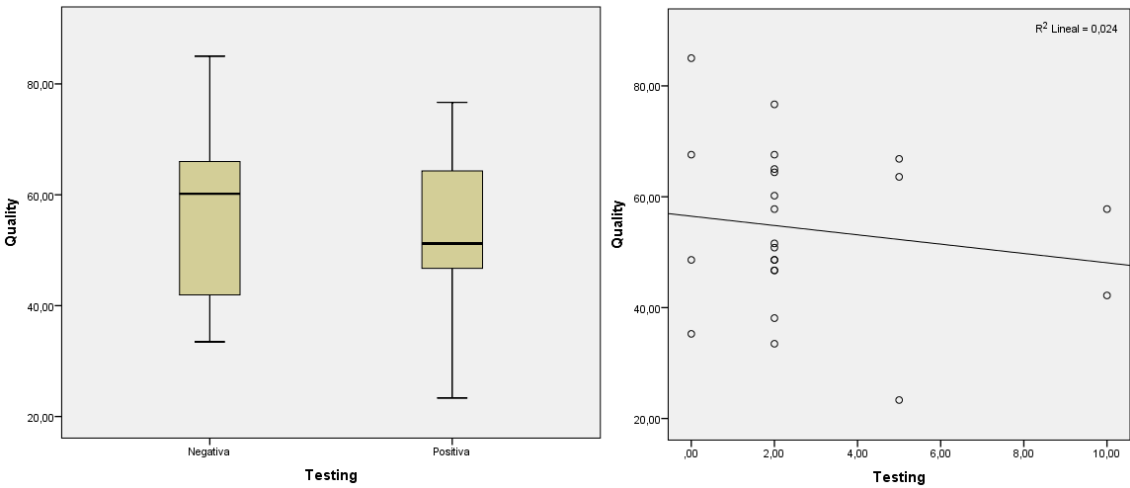
3.10 CALIDAD – EXPERIENCIA DE PROGRAMACIÓN



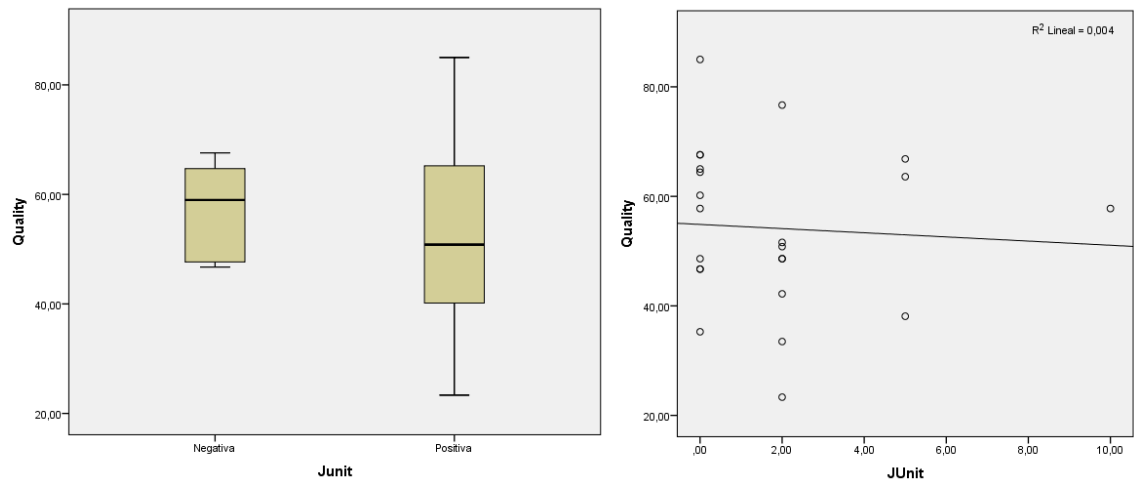
3.11 CALIDAD – EXPERIENCIA EN JAVA



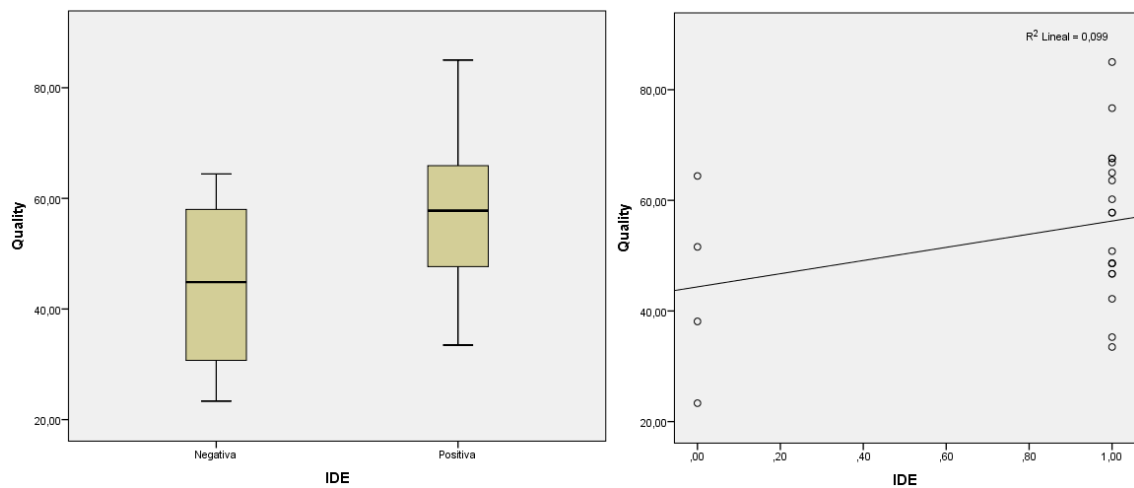
3.12 CALIDAD – EXPERIENCIA EN TESTING



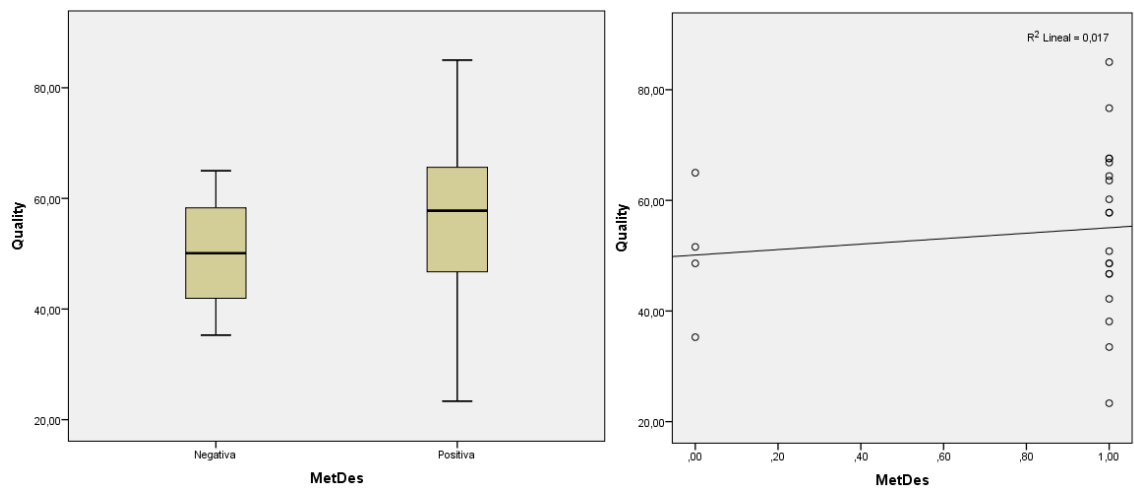
3.13 CALIDAD – EXPERIENCIA EN JUNIT



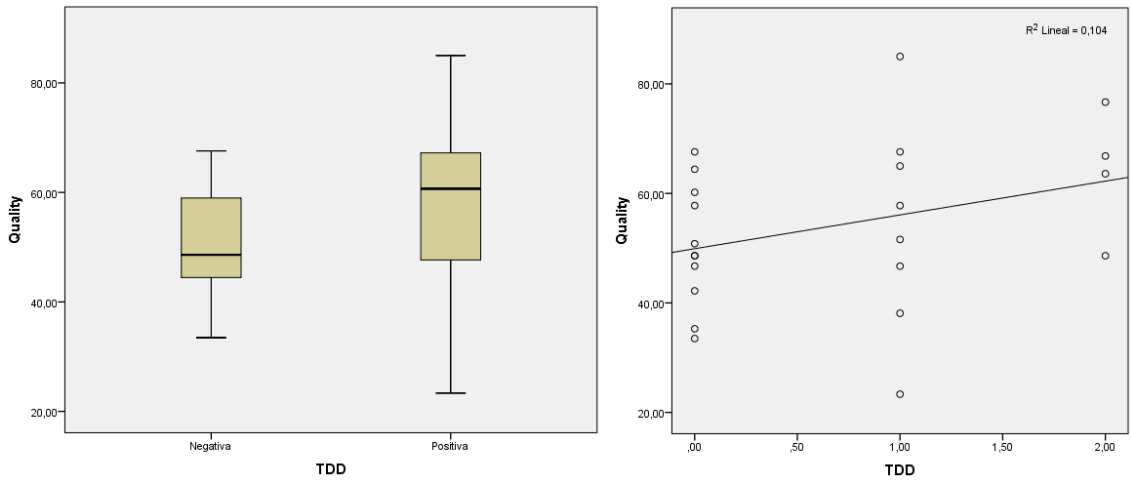
3.14 CALIDAD – EXPERIENCIA IDE



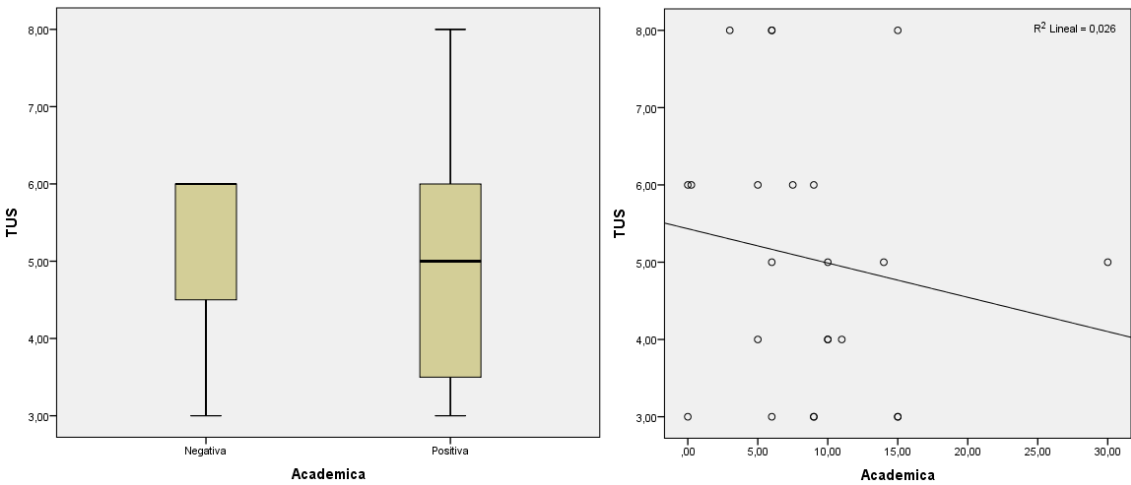
3.15 CALIDAD – EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO



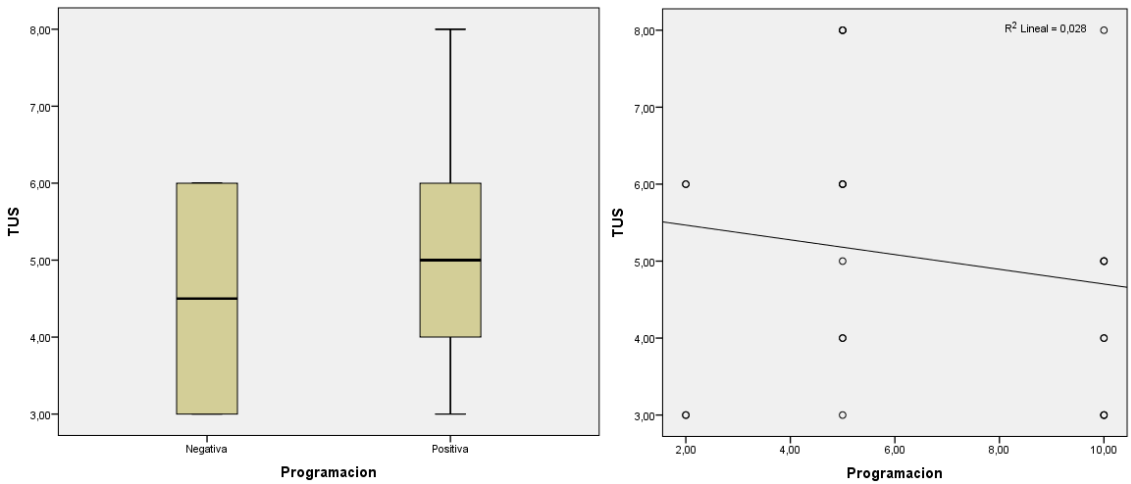
3.16 CALIDAD – EXPERIENCIA EN TDD



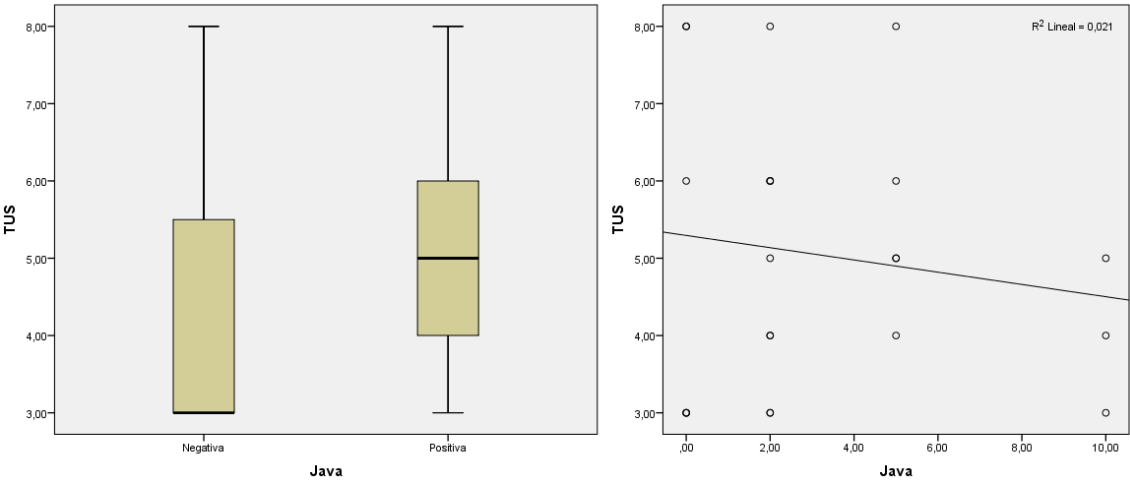
3.17 TUS – EXPERIENCIA ACADÉMICA



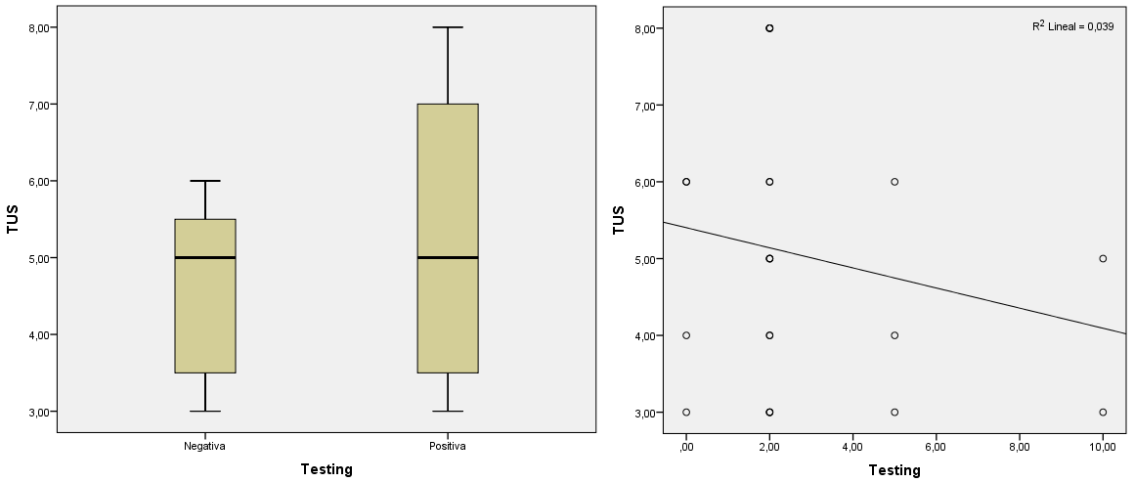
3.18 TUS – EXPERIENCIA DE PROGRAMACIÓN



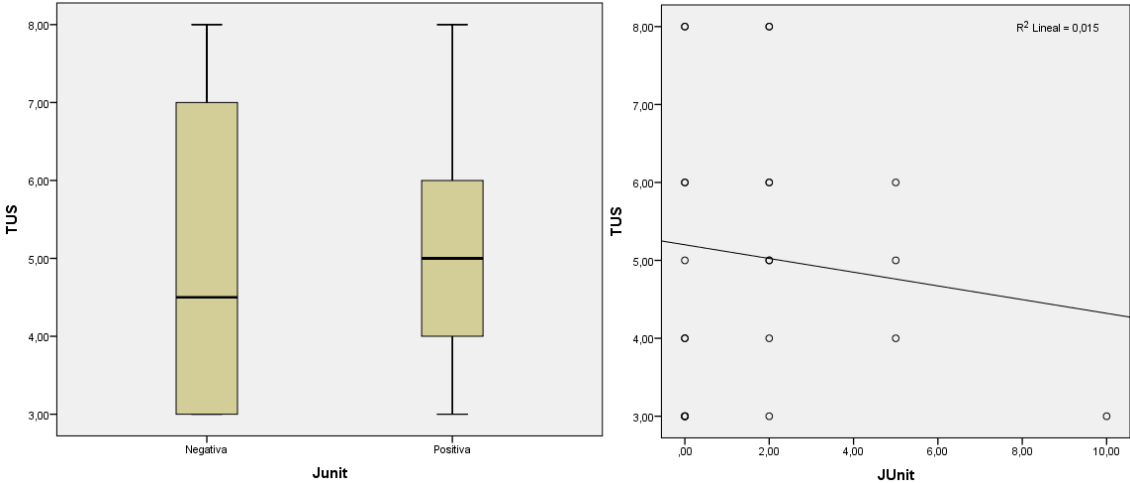
3.19 TUS – EXPERIENCIA EN JAVA



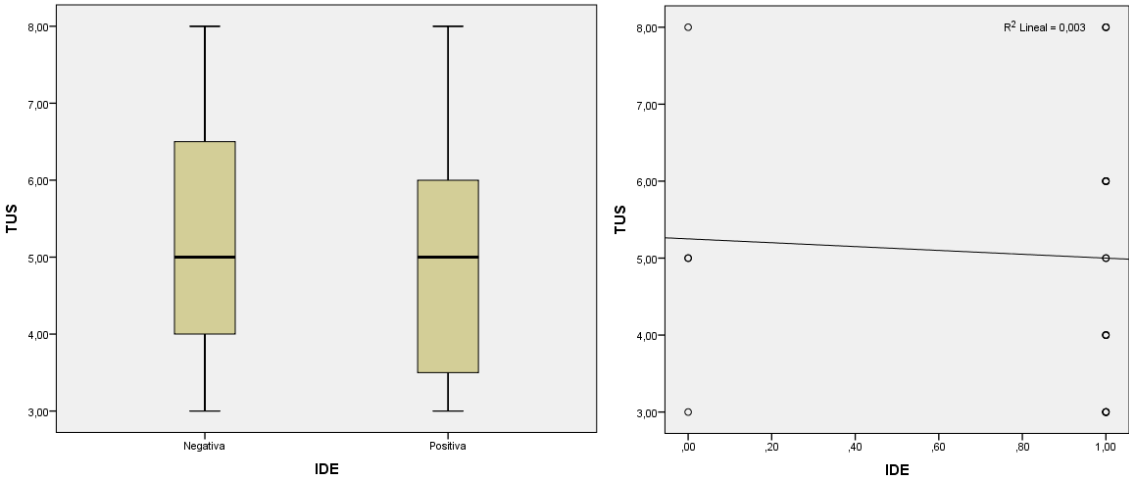
3.20 TUS – EXPERIENCIA EN TESTING



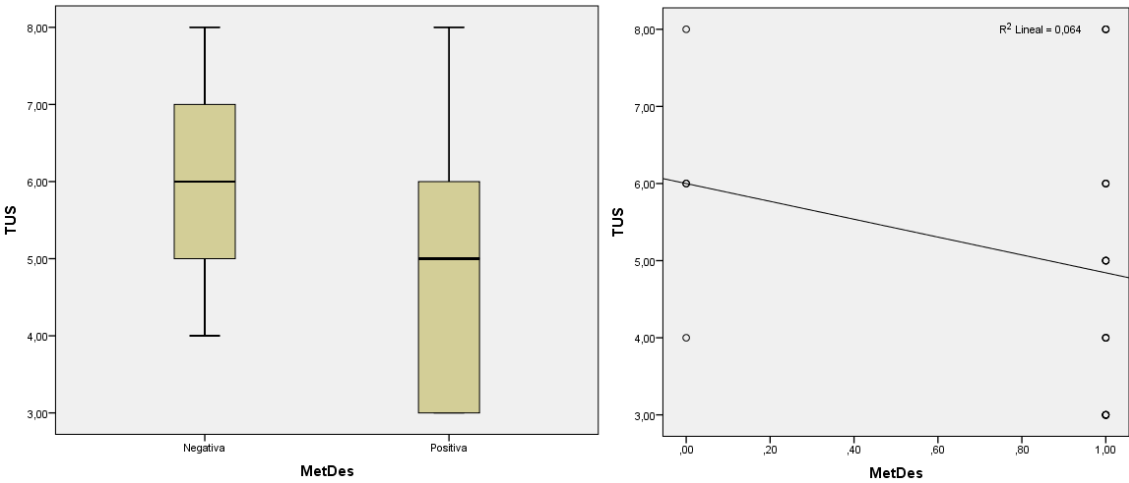
3.21 TUS – EXPERIENCIA EN JUNIT



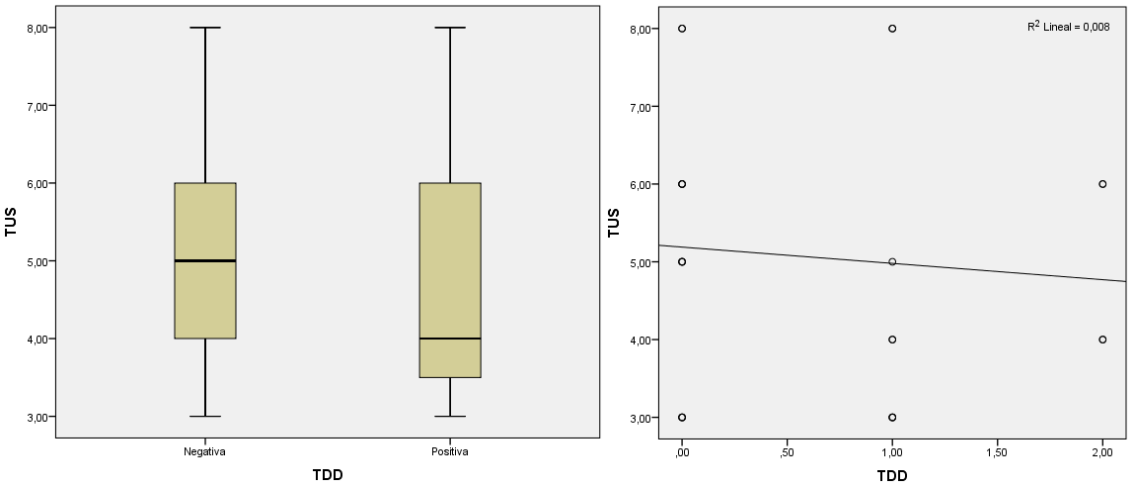
3.22 TUS – EXPERIENCIA IDE



3.23 TUS – EXPERIENCIA EN METODOLOGÍAS DE DESARROLLO



3.24 TUS – EXPERIENCIA EN TDD



ANEXO E.
Formulario Demográfico

Demographics

Pre-questionnaire

* Required

Respondent ID *

Company *

Location *

Function *

1. Please indicate the title(s) and degree(s) of your education (e.g., BS in computer science, MS in management). *

2. Please indicate the roles that you have worked during your professional career. (e.g., developer XX months/years, tester YY months/years, etc.). *

3. Please indicate the programming languages you have used (including your education), and the number of years of experience with each one.

Programming Language 1

3.1.1 Programming language *

3.1.2 Years (Education) *

3.1.3 Years (Professional career) *

Programming Language 2

3.2.1 Programming language

3.2.2 Years (Education)

3.2.3 Years (Professional career)

Programming Language 3

3.3.1 Programming language

3.3.2 Years (Education)

3.3.3 Years (Professional career)

4. How would you rate your experience with programming? *

- ☐ No experience (<2 years)
- ☐ Novice (2-<=5 years)
- ☐ Intermediate (5-<=10 years)
- ☐ Expert (>10 years)

5. Which development methodologies have you used so far? (e.g., waterfall, iterative, spiral, agile. If you choose agile, please indicate the type (scrum, tdd, xp, etc.). Include the methodologies you used in academia as well. Indicate the number of years of experience with each one (e.g., waterfall, 1 year of education, 5 years of professional career). 5.1.1

Methodology *

5.1.2 Years (Education) *

5.1.3 Years (Professional career) *

5.2.1 Methodology

5.2.2 Years (Education)

5.2.3 Years (Professional career)

5.3.1 Methodology

5.3.2 Years (Education)

5.3.3 Years (Professional career)

6. How would you rate your experience in unit testing? *

- ☐ No experience (<2 years)
- ☐ Novice (2-<=5 years)
- ☐ Intermediate (5-<=10 years)
- ☐ Expert (>10 years)

7. Are you currently using a tool for unit testing (for executing, monitoring)? If yes, please write the names of the tools. *

8. What IDE (Integrated Development Environment) do you currently use? *

9. How would you rate your experience in Java? *

- ☐ No experience (<2 years)
- ☐ Novice (2-<=5 years)

- ☐ Intermediate (5-<=10 years)
- ☐ Expert (>10 years)

10. How would you rate your experience with Junit testing framework? *

- ☐ No experience (<2 years)
- ☐ Novice (2-<=5 years)
- ☐ Intermediate (5-<=10 years)
- ☐ Expert (>10 years)

11. Have you ever used TDD as a development methodology? *

- ☐ Yes
- ☐ No

12. If “Yes” above, how would you rate your experience with TDD?

- ☐ No experience (<2 years)
- ☐ Novice (2-<=5 years)
- ☐ Intermediate (5-<=10 years)
- ☐ Expert (>10 years)

13. Please indicate the certification(s) you have received during your professional career. (e.g., SEI certification as Personal Software Process (PSP) developer, CMMI certification, or ITIL certification as application engineer) *

14. Have you ever attended any training on testing, or more specifically unit testing? If yes, please briefly explain its content. *

15. Have you ever attended any training on TDD? *

- ☐ Yes
- ☐ No

16. If “Yes” please briefly explain with regard to the following questions:

a) What was taught during the training?

b) How long did the training take? (in days or hours if possible)

c) When did you take the training?

d) Did you take the training at your current occupation?

e) Are you still practicing TDD? Why?

17. Have you ever been involved in TDD studies in industry? *

☐ Yes

☐ No

18. If “Yes” above, please share the results you got from the pilot study, and your opinion on its effectiveness.

19. Have you ever attended any coding kata? *

- ☐ Yes
- ☐ No

19a. If "yes", please indicate when and which katas (name of the coding task) you completed

Submit

Never submit passwords through Google Forms.